

On $\sigma(n)$ and $\varphi(n)$

by

PAN CHENG-TUNG

Presented by W. SIERPIŃSKI on July 30, 1956

The aim of this note is to give some improvements on several results, previously obtained by Walfisz [1], Davenport and Chowla, concerning the average or the Riesz mean of the arithmetical function $\sigma(n)$, $\frac{\sigma(n)}{n}$, $\varphi(n)$ or $\nu_4(n)$. Here $\sigma(n)$ is the sum of the divisors of n , and $\nu_4(n)$ is the number of representations of n as a sum of four squares of integers. The main tool is the use of Vinogradov's new method of estimating trigonometrical sums.

Let

$$\sum_{n \leq x} \sigma(n) = \frac{1}{12} \pi^2 x^2 + \Delta(x),$$

$$\sum_{n \leq x} \varphi(n) = \frac{3}{\pi^2} x^2 + R(x),$$

and

$$\sum_{n \leq x} \nu_4(n) = \frac{\pi^2}{2} x^2 + T(x).$$

The main results obtained in this paper are as follows.

THEOREM 1. $\Delta(x) = O\{x(\log x \log \log x)^{3/4}\}$.

This is better than the result

$$\Delta(x) = O(x \log^{4/5+\varepsilon} x)$$

which was obtained independently by Walfisz [1] and Davenport [2].

From this we deduce that

THEOREM 2. $T(x) = O\{x(\log x \log \log x)^{3/4}\}$.

THEOREM 3. Let m denote a positive integer, then we have

$$\sum_{n \leq m} \frac{\sigma(n)}{n} \log \frac{m}{n} = \frac{\pi^2}{6} m - \frac{1}{4} \log^2 m + a \log m + b + O(m^{-2/3+\varepsilon}),$$

where a and b are absolute constants, and ε is an arbitrary positive number.

THEOREM 4. Let x be a positive real number, then

$$\sum_{n \leq x} \frac{R(n)}{n} = \frac{3}{\pi^2} x + O\{x \exp(-A \log^{4/7-\varepsilon} x)\},$$

where A is a positive number, and ε is an arbitrarily small positive number.

This is an improvement on Chowla's [3] result

$$\sum_{n \leq x} R(n) \sim \frac{3}{2\pi^2} x^2.$$

From this we obtain

THEOREM 5.

$$\sum_{n \leq x} \varphi(n) \log \frac{x}{n} = \frac{3}{2\pi^2} x^2 - \frac{3}{\pi^2} \{x\} x + O(x \exp(-A \log^{4/7-\varepsilon} x)),$$

where $\{x\}$ represents the fractional part of x and A is a positive constant.

The proofs of the Theorems will be published in Chinese in *Acta Scientiarum Naturalium Universitatis Pekinensis*.

This paper has been written under Professor Min. Szu-Hoa's supervision.

DEPARTMENT OF MATHEMATICS AND MECHANICS, PEKING UNIVERSITY

REFERENCES

- [1] A. Walfisz, *Teilerprobleme*, Math. Zeits. **26** (1927), 66-88.
- [2] H. Davenport, *A divisor problem*, Quart. J. of Math. (Oxford) **20** (1949).
- [3] S. S. Pillai and S. P. Chowla, *On the error term in some asymptotic formulae on the theory of numbers*, J. London Math. Soc., **5** (1930), 95-101.

Fractional Integrals of Functions of Several Variables with Applications to the Theory of Approximation

by

CHENG MIN-TEH and CHEN YUNG-HOO

Presented by W. ORLICZ on June 14, 1956

Let $F(P)$ be a continuous function in the m -dimensional Euclidean space Ω , and let the integral

$$\int_{\Omega} F(Q) r_{PQ}^{a-m} dQ$$

be convergent absolutely for $0 < a < m$, where r_{PQ} denotes the distance between P and Q . Riesz [2] defined the fractional integral of $F(P)$ of order a as the integral

$$I_a[F(P)] = \frac{1}{H_m(\alpha)} \int_{\Omega} F(Q) r_{PQ}^{a-m} dQ \quad (0 < a < m),$$

where

$$H_m(\alpha) = e^{\frac{\pi}{2} i \alpha} 2^{\frac{\alpha}{2}} \pi^{\frac{m-1}{2}} \frac{\Gamma\left(\frac{\alpha}{2}\right)}{\Gamma\left(\frac{m-\alpha}{2}\right)}.$$

Riesz's definition is not entirely satisfactory for periodic functions. Our main purpose is to extend Riesz's definition to the periodic case.

We assume throughout this paper that $F(P)$ is a continuous function with period 1 for each variable and with mean value zero, i. e.

$$\int_0^1 \dots \int_0^1 F(P) dP = 0.$$

Let the Fourier series of $F(P)$ be

$$\sum_{-\infty}^{\infty} C_{n_1, \dots, n_m} e^{2\pi i(n_1 x_1 + \dots + n_m x_m)}.$$

Here $\sum'$ denotes the summation, where the term corresponding to $n_1 = n_2 = \dots = n_m = 0$ is absent. Let

$$A_\nu(P) = \sum_{n_1^2 + \dots + n_m^2 = \nu} C_{n_1, \dots, n_m} e^{2\pi i(n_1 x_1 + \dots + n_m x_m)}.$$

Then we define

$$\lim_{R \rightarrow \infty} \sum'_{\nu \leq R^2} \left(1 - \frac{\nu}{R^2}\right)^{\sigma_m} \frac{A_\nu(P)}{(2\pi)^{\alpha} \nu^{\frac{1}{2}\alpha} e^{\frac{\pi}{2} i \alpha}} = I_\alpha^*[f(P)] \left(\sigma_m = \left[\frac{m-1}{2}\right] + 1, \alpha \geq 0\right)$$

as the fractional integral of $F(P)$ of order α . The left-hand side of the above expression is uniformly convergent.

Concerning the fractional integration we have the following theorems:

THEOREM 1. For $0 < \alpha < 1$, we have

$$I_\alpha^*[F(P)] = I_\alpha[F(P)];$$

moreover, the following two relations hold:

$$\begin{aligned} I_\alpha^*[I_\beta^*F(P)] &= I_{\alpha+\beta}^*[F(P)] & (\alpha \geq 0, \beta \geq 0), \\ \Delta I_{\alpha+2}^*[F(P)] &= I_\alpha^*[F(P)] & (x \geq 0). \end{aligned}$$

Let us define the mean modulus of continuity $\omega^*(\varrho)$ of $F(P)$ as:

$$\omega^*(\varrho) = \max_{\substack{P \in D \\ 0 \leq t \leq \varrho}} \left| \frac{1}{S(t)} \int_{S(P; t)} \{F(R) - F(P)\} dR \right|,$$

where $S(t)$ is the surface volume of the sphere $S(P; t)$. If $\omega^*(\varrho) = O(\varrho^\alpha)$, we say that $F \in \text{Lip}^* \alpha$. It is easy to see that the class $\text{Lip}^* \alpha$ is equivalent to the ordinary Lipschitz class $\text{Lip} \alpha$ for $0 < \alpha < 1$.

THEOREM 2. If $F(P) \in \text{Lip} \alpha$ ($0 < \alpha \leq 1$), then for $0 < \beta < 1$, we have

$$I_\beta^*[F(P)] \in \text{Lip}^*(\alpha + \beta).$$

In particular, if $\alpha + \beta < 1$, then from the assumption $F(P) \in \text{Lip} \alpha$, we obtain

$$I_\beta^*[F(P)] \in \text{Lip}(\alpha + \beta).$$

Now let us define the fractional Laplace operator Δ^β of order β ($0 < \beta < 1$) as follows:

$$\Delta^\beta F(P) = \Delta I_{2(1-\beta)}^*[F(P)],$$

whenever the right-hand side exists.

THEOREM 3. If $F(P) \in \text{Lip}^* \alpha$, then for $1 < 2\beta < \alpha \leq 2$, we have

$$\Delta^\beta F(P) \in \text{Lip}^*(\alpha - 2\beta).$$

In virtue of the uniqueness theorem of double trigonometrical series [2], we establish the following

THEOREM 4. If $\Delta^\beta F(x, y) = G(x, y)$ ($0 < \beta < 1$) is continuous, then

$$I_{2\beta}^*[G(x, y)] = F(x, y).$$

The remaining part of our paper deals with the approximation of functions of two variables by trigonometrical polynomials.

Suppose that $f(x, y)$ is periodic with period 2π and belongs to $C^{2p}(D)$, i. e. $f(x, y)$ has continuous $2p$ -th partial derivatives on the torus D : $0 \leq x \leq 2\pi$, $0 \leq y \leq 2\pi$. Let $\bar{\omega}_{2p}(\varrho)$ be the modulus of continuity of $\Delta^p f(x, y)$, where Δ^p denotes the ordinary iterated Laplace operator of order p . We write

$$S_R^{k,\delta}(x, y; f) = \sum_{v \leq R^2} \left(1 - \frac{v^{k/2}}{R^k}\right)^\delta A_v(x, y),$$

where

$$A_v(x, y) = \sum_{m^2 + n^2 = v} C_{m,n} e^{i(mx + ny)},$$

$C_{m,n}$ being the Fourier coefficients of $f(x, y)$. Then we have

THEOREM 5. *If $f(x, y) \in C^{2p}(D)$, then for $k > 2p$, $\delta > 2p + \frac{1}{2}$,*

$$S_R^{k,\delta}(x, y; f) - f(x, y) = O\left(\frac{1}{R^{2p}} \bar{\omega}_{2p}\left(\frac{1}{R}\right)\right)$$

holds uniformly on D .

The above theorem gives the relation between the degree of approximation and the modulus of continuity $\bar{\omega}_{2p}(\varrho)$ of $\Delta^p f(x, y)$, which is not related to all the $2p$ -th partial derivatives of $f(x, y)$.

From Theorems 4 and 5, we obtain

THEOREM 6. *Let $\bar{\omega}_{2\beta}(\varrho)$ be the modulus of continuity of $\Delta^\beta F(x, y)$ with $0 < \beta < \frac{1}{2}$. If*

$$\int_0^1 \frac{\bar{\omega}_{2\beta}(\varrho)}{\varrho} d\varrho < \infty,$$

then uniformly on D

$$S_R^{k,\delta}(x, y; F) - F(x, y) = O\left(\frac{1}{R^{2\beta}} \bar{\omega}_{2\beta}\left(\frac{1}{R}\right)\right),$$

where

$$S_R^{k,\delta}(x, y; F) = \sum'_{v \leq R^2} \left(1 - \frac{v^{k/2}}{R^k}\right)^\delta A_v(x, y)$$

with $k > 2$ and $\delta > \frac{7}{2}$, $A_v(x, y)$ being the corresponding expression with respect to $F(x, y)$.

The proofs will appear in Science Report of Peking University.

PEKING UNIVERSITY

REFERENCES

- [1] M. T. Cheng, *Uniqueness of multiple trigonometric series*, Ann. Math. **52** (1950), 403.
- [2] M. Riesz, *L'intégrale de Riemann-Liouville et le problème de Cauchy*, Acta Math. **81** (1949), 1.

A Generalisation of Mikusiński's Operational Calculus

by

W. SŁOWIKOWSKI

Presented by W. ORLICZ on June 14, 1956

1. Introduction. This paper treats of the differential equation

$$(1) \quad x^{(n)}(t) + A_1 x^{(n-1)}(t) + \dots + A_n x(t) = a(t),$$

where $x(t)$, $y(t)$ are continuous functions of the real variable t with values in a linear topological space X , and $A_1, \dots, A_n$ are continuous endomorphisms of X . We shall prove the following

THEOREM. *Let X be locally convex and sequentially complete, the topology in X being determined by a set of pseudonorms $|x|_\lambda$ ($\lambda \in A$).*

Let $x_0, x_1, \dots, x_{n-1} \in X$. If there exist positive integers $n_1, n_2, \dots, n_k$, $n_1 + n_2 + \dots + n_k = n$ and endomorphisms $B_1, B_2, \dots, B_k$ such that

$$(i) \quad B_i B_j = B_j B_i \quad (i, j = 1, 2, \dots, k),$$

(ii) *the B_i are continuous with respect to each pseudonorm separately (i. e., for each $\lambda \in A$ and $1 \leq i \leq k$ there is a number $M_{i\lambda}$ such that $|B_i x|_\lambda \leq M_{i\lambda} |x|_\lambda$ for $x \in X$)*

$$(iii) \quad \xi^n + A_1 \xi^{n-1} + \dots + A_n \xi = (\xi - B_1)^{n_1} (\xi - B_2)^{n_2} \dots (\xi - B_k)^{n_k}$$

formally, then equation (1) has exactly one solution $x(t)$ such that

$$x^{(j)}(0) = x_j \quad (j = 1, 2, \dots, n-1).$$

We shall also give a simple formula for the solution $x(t)$ of (1) (see formula (12) below).

Some applications of the theorem will be given in a separate paper. The solution of (1) is found by a method which is a generalisation of Mikusiński's operational calculus [1]. This generalisation is the subject of 2.

It will be convenient to consider the operators $A_1, \dots, A_n, B_1, \dots, B_k$ as the elements of a sequentially complete commutative algebra R of endomorphisms of X continuous with respect to each pseudonorm separately and containing the identity I .

The algebra R is topologised by pseudonorms

$$(2) \quad |A|_\lambda \stackrel{\text{def}}{=} \sup_{|x|_\lambda \leq 1} |A_x|_\lambda \quad (A \in R).$$

For instance R can be defined as the least sequentially complete algebra containing $B_1, B_2, \dots, B_k, I$.

2. Generalisation of Mikusiński's operational calculus [1].

In what follows R is a locally convex, sequentially complete commutative algebra of endomorphisms of the space X , continuous with respect to each pseudonorm separately. R is topologised by (2) and contains the identity I .

We denote by $\mathbf{X}$ the linear space of all continuous functions $\mathbf{x} = \{x(t)\}$ of real variables with values in X , and by $\mathbf{R}$ the linear space of all continuous functions $\mathbf{A} = \{A(t)\}$ of real variables with values in R . We shall consider the following two types of convolutions of such functions:

$$(3) \quad \mathbf{A}\mathbf{x} = \left\{ \int_0^t A(t-\tau)x(\tau) d\tau \right\} \in \mathbf{X},$$

$$(4) \quad \mathbf{A}\mathbf{B} = \left\{ \int_0^t A(t-\tau)B(\tau) d\tau \right\} \in \mathbf{R},$$

where $\mathbf{A} = \{A(t)\}$, $\mathbf{B} = \{B(t)\} \in \mathbf{R}$ and $\mathbf{x} = \{x(t)\} \in \mathbf{X}$.

It can be shown that

$$(5) \quad \begin{aligned} \mathbf{A}(\mathbf{B}\mathbf{x}) &= (\mathbf{A}\mathbf{B})\mathbf{x}; & (\mathbf{A}\mathbf{B})\mathbf{C} &= \mathbf{A}(\mathbf{B}\mathbf{C}) & \text{and} & & \mathbf{A}\mathbf{B} &= \mathbf{B}\mathbf{A} \\ & \text{for } \mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathbf{R} & \text{and} & & \mathbf{x} \in \mathbf{X}. \end{aligned}$$

The linear space $\mathbf{R}$ with convolution (4) defining multiplication is a linear ring of endomorphisms of $\mathbf{X}$. The result of the application of $\mathbf{A} \in \mathbf{R}$ to the element $\mathbf{x} \in \mathbf{X}$ is the element $\mathbf{A}\mathbf{x} \in \mathbf{X}$ defined in (3).

Besides (3) and (4) we shall also use the following products:

$$(6) \quad \mathbf{A}\mathbf{x} = \{Ax(t)\} \in \mathbf{X} \quad \left| \quad \text{where } \mathbf{x} = \{x(t)\} \in \mathbf{X}, \right.$$

$$(7) \quad \mathbf{A}\mathbf{B} = \{AB(t)\} \in \mathbf{R} \quad \left| \quad \mathbf{A} = \{A(t)\}, \quad \mathbf{B} = \{B(t)\} \in \mathbf{R}, \right.$$

$$(8) \quad \mathbf{A}\mathbf{x} = \{A(t)x\} \in \mathbf{X} \quad \left| \quad \text{and } A \in R; \quad x \in X. \right.$$

The following equalities can be proved:

$$(9) \quad \begin{aligned} \mathbf{A}(\mathbf{B}\mathbf{x}) &= (\mathbf{A}\mathbf{B})\mathbf{x}; & \mathbf{A}(\mathbf{B}\mathbf{x}) &= (\mathbf{A}\mathbf{B})\mathbf{x}; & (\mathbf{A}\mathbf{B})\mathbf{C} &= \mathbf{A}(\mathbf{B}\mathbf{C}) \\ & \mathbf{A}(\mathbf{B}\mathbf{x}) &= (\mathbf{A}\mathbf{B})\mathbf{x}; & \mathbf{A}(\mathbf{B}\mathbf{x}) &= (\mathbf{A}\mathbf{B})\mathbf{x}. \end{aligned}$$

The function $\mathbf{L} = \{L(t)\}$, where $L(t) = I$ for every t , is the integration operator in $\mathbf{X}$: $\mathbf{L}\mathbf{x} = \left\{ \int_0^t x(\tau) d\tau \right\}$. Hence, $\mathbf{L}\mathbf{x} = \mathbf{0}$ implies $\mathbf{x} = \mathbf{0}$.

We denote by $\mathbf{S}$ the operation $\mathbf{L}^{-1}$ defined on the linear subspace $\mathbf{L}(\mathbf{X}) \subset \mathbf{X}$. It follows from [2] that there exists exactly one (up to isomorphism) linear space $\hat{\mathbf{X}} \supset \mathbf{X}$ such that the operation $\mathbf{S}$ can be extended to an automorphism of $\hat{\mathbf{X}}$ onto $\hat{\mathbf{X}}$ (denoted by the same letter $\mathbf{S}$), and

that every element $\hat{x} \in \hat{X}$ has the form

$$\hat{x} = S^n x, \quad \text{where } x \in X$$

and S^n is the n -th iteration of S .

Every endomorphism $A \in R$ can be extended to an endomorphism of $\hat{X}$ as follows:

$$A\hat{x} = S^n(Ax), \quad \text{where } \hat{x} = S^n x \quad \text{and} \quad x \in X.$$

Consequently, we may consider R as an algebra of endomorphisms of $\hat{X}$. The multiplications (7) may also be extended:

$$A\hat{x} = S^n(Ax), \quad \text{where } A \in R; \quad \hat{x} = S^n x \in \hat{X} \quad \text{and} \quad x \in X.$$

Let $\hat{R}$ be the least algebra of endomorphisms of $\hat{X}$ containing R and the automorphism S . All elements of $\hat{R}$ may be represented in the form $S^n A$, where $A \in R$ and S^n is the n -th iteration of S . The element $I = LS$ is the identity of $\hat{R}$. There is no topology defined in $\hat{R}$.

The multiplications (7) and (8) can be extended on $\hat{R}$ as follows:

$$Ax = S^n(Bx),$$

$$CA = S^n(CB), \quad \text{where } A = S^n B, \quad B \in R, \quad x \in X \quad \text{and} \quad C \in R.$$

Equations (5) and (9) are also valid for arbitrary elements of $\hat{X}$ and $\hat{R}$.

The linear space X and the algebra R can be isomorphically imbedded in the space $\hat{X}$ and the algebra $\hat{R}$ respectively as follows:

$$x \mapsto Ix,$$

$$A \mapsto IA.$$

Consequently, we shall consider X and R as subsets of $\hat{X}$ and $\hat{R}$, respectively identifying x with Ix and A with IA .

If $x = \{x(t)\} \in X$ and $x' = \{x'(t)\} \in X$, then

$$Lx' = \left\{ \int_0^t x'(\tau) d\tau \right\} = \{x(t) - x(0)\} = x - \{x(0)\} = x - Lx(0).$$

Multiplying both sides of this equation by S we obtain $x' = Sx - x(0)$. Hence by induction we obtain the formula for the l -th derivative $x^{(l)}$ of $x = \{x(t)\} \in X$:

$$x^{(l)} = S^l x - (x^{(l-1)}(0) + Sx^{(l-2)}(0) + \dots + S^{l-2}x'(0) + S^{l-1}x(0)).$$

3. Proof of the Theorem. Equation (1) can now be written in the form

$$(S^n + A_1 S^{n-1} + \dots + A_{n-1} S + A_n)x = (S^{n-1}y_0 + S^{n-2}y_1 + \dots + Sy_{n-2} + y_{n-1}) + a,$$

or by (iii) in the form

$$(10) \quad (S - B_1)^{n_1} (S - B_2)^{n_2} \dots (S - B_k)^{n_k} x = S^{n-1}y_0 + \dots + Sy_{n-2} + y_{n-1} + a,$$

where $y_i = x^{(i)}(0) + A_1 x^{(i-1)}(0) + \dots + A_i x(0)$ and $\mathbf{a} = \{a(t)\}$.

The series $e^B = \sum_{m=0}^{\infty} \frac{B^m}{m!}$ converges with respect to each pseudonorm (2).

Its sum e^B is an element of R . Hence $\{e^{tB}\} \in \mathbf{R} \subset \hat{\mathbf{R}}$. It is easy to verify that for every $B \in R$ the element $(S-B)^{-1}$ exists. Moreover,

$$(11) \quad (S-B)^{-l} = \left\{ \frac{t^{l-1}}{(n-1)!} e^{Bt} \right\} \quad \text{for } l=1, 2, \dots$$

The element $(S-B_1)^{n_1} \dots (S-B_k)^{n_k}$ also has an inverse. Consequently equation (10) has exactly one solution:

$$(12) \quad \mathbf{x} = \frac{S^{n-1} y_0 + \dots + S y_{n-2} + y_{n-1} + \mathbf{a}}{(S-B_1)^{n_1} (S-B_2)^{n_2} \dots (S-B_k)^{n_k}} \in \hat{X}.$$

We shall show that this solution $\mathbf{x}$ always belongs to X . This follows from (11) and the fact that the operator

$$\frac{S^l}{(S-B_1)^{n_1} (S-B_2)^{n_2} \dots (S-B_k)^{n_k}} \quad l \leq n$$

is a sum of operators of the form

$$\frac{C}{(S-B)^{l_1} \dots (S-B)^{k_1}},$$

where $l_i \leq n_i$. The proof is by induction based on the following identity:

$$\frac{S^l}{W(S)} = \frac{S^l - B_i^l}{W(S)} = \frac{B_i^l}{W(S)} = \frac{\xi^{l-1} + B_i S^{l-2} + \dots + B_i^{l-2} S + B_i^{l-1}}{W_i(S)} + \frac{B_i^l}{W(S)},$$

where $W(S) = (S-B_1)^{l_1} \dots (S-B_k)^{l_k} l + \dots + l_k \leq l$; $l_i \geq 1$ and $W_i(S) = (S-B_1)^{l_1} \dots (S-B_i)^{l_i-1} \dots (S-B_k)^{l_k}$.

Notice that in the special case, where the elements $B_i - B_j$ $i \neq j$ have inverses, the quotient

$$\frac{C}{(S-B_1)^{l_1} (S-B_2)^{l_2} \dots (S-B_k)^{l_k}}, \quad \text{where } C \in R,$$

may be represented as a linear combination of simple quotients of the form $D/(S-B_i)^m$, where $D \in R$. The inductive proof is based on the following identity:

$$\frac{I}{(S-A)(S-B)} = (A-B)^{-1} \left\{ \frac{I}{(S-A)} - \frac{I}{(S-B)} \right\}.$$

It can be seen in the preceding proof that only commutativity of L and elements of R has been necessary. Then, after a few modifications, the proposed calculus is valid for arbitrary endomorphisms (not necessary commutative).

4. EXAMPLE. Let us consider the Hilbert space H ($\|x\|$) and the self-adjoint operator A with continuous inversion.

Then there are projections P_n commutative with A and such that $\sum_{n=1}^{\infty} P_n = I$, $P_n P_m = 0$ for $m \neq n$ and AP_n are defined on the whole H and continuous. Let $\|x\|_n = \|P_n x\|$ and let X be an extension of $H(\|x\|_n)$ to complete space. Operator A can be extended to the whole X and satisfy the condition (ii) of the Theorem.

Further we have:

$$\frac{I}{S^2 + A^2} = \{A^{-1} \sin tA\}$$

$$\frac{S - A}{S^2 + A^2} = \{\cos tA\}.$$

We can now consider the equation

$$x''(t) + A^2 x(t) = a(t).$$

Its solution is

$$\{x(t)\} = \frac{Sx(0) + x'(0) + a}{S^2 + A^2} = \frac{(S - A)x(0)}{S^2 + A^2} + \frac{Ax(0) + x'(0)}{S^2 + A^2} + \frac{a}{S^2 + A^2}.$$

Then

$$x(t) = (\cos tA + \sin tA)x(0) + (A^{-1} \sin tA)x'(0) + \int_0^t (A^{-1} \sin(t-\tau)A)a(\tau) d(\tau).$$

If $A^2 = \int_0^{\infty} \lambda dE_{\lambda}$ is the spectral representation of A^2 and $P_n = E_n - E_{n-1}$, then we have

$$x(t) = \int_0^{\infty} \cos t \sqrt{\lambda} + \sin t \sqrt{\lambda} dE_{\lambda} x(0) + \int_0^{\infty} \frac{\sin t \sqrt{\lambda}}{\sqrt{\lambda}} dE_{\lambda} x'(0) +$$

$$+ \int_0^t \left(\int_0^{\infty} \frac{\sin \sqrt{\lambda}(t-\tau)}{\sqrt{\lambda}} dE_{\lambda}(\tau) \right) d\tau.$$

If $-A^2$ is an elliptic operator in the usual sense, then $x(t)$ is the solution of wave equation.

INSTITUTE OF MATHEMATICS, POLISH ACADEMY OF SCIENCES

REFERENCES

- [1] J. G. Mikusiński, *Sur les fondements du calcul opératoire*, *Studia Mathematica* **11** (1950), 41.
- [2] W. Słowikowski, *A generalisation of the theory of distributions*, *Bull. Acad. Polon. Sci. Cl. III.* **3** (1955), 3.

A Theorem on Non-Classical Functional Calculi

by

R. SIKORSKI

Presented by A. MOSTOWSKI on July 23, 1956

This paper is a supplement to paper [1], the same notation and terminology being used. We recall that if $\mathcal{X}$ is a topological space, then $\mathbf{C}(\mathcal{X})$ denotes the closure algebra of all subsets of $\mathcal{X}$, and $\mathbf{H}(\mathcal{X})$ — the algebra of all open subsets of $\mathcal{X}$ ([1], p. 85). If a is a formula from the Lewis functional calculus $\mathcal{S}_\lambda^*$, (or from the Heyting functional calculus $\mathcal{S}_\pi^*$, or from the positive functional calculus $\mathcal{S}_\pi^*$), then $(I_0, \mathbf{C}(\mathcal{X}))\Phi_a$ (then $(I_0, \mathbf{H}(\mathcal{X}))\Phi_a$) denotes the functional obtained from a by interpreting the individual variables as variables running over the set I_0 of positive integers, and the k -argument functional variables as k -argument functions from I_0 into $\mathbf{C}(\mathcal{X})$ (into $\mathbf{H}(\mathcal{X})$) — see [1] p. 71. We write $(I_0, \mathbf{C}(\mathcal{X}))\Phi_a = \mathcal{X}$ (or $(I_0, \mathbf{H}(\mathcal{X}))\Phi_a = \mathcal{X}$) if this functional is identically equal to $\mathcal{X}$.

It was proved in [1] (pp. 83, 87 and 89) that there exists a topological space $\mathcal{X}_\lambda (\mathcal{X}_\pi, \mathcal{X}_\pi)$ such that, for every formula a from the Lewis functional calculus $\mathcal{S}_\lambda^*$ (from the Heyting functional calculus $\mathcal{S}_\pi^*$, from the positive functional calculus $\mathcal{S}_\pi^*$), a is provable in $\mathcal{S}_\lambda^*$ (in $\mathcal{S}_\pi^*$, in $\mathcal{S}_\pi^*$) if, and only if, $(I_0, \mathbf{C}(\mathcal{X}_\lambda))\Phi_a = \mathcal{X}_\lambda$ (if $(I_0, \mathbf{H}(\mathcal{X}_\pi))\Phi_a = \mathcal{X}_\pi$, if $(I_0, \mathbf{H}(\mathcal{X}_\pi))\Phi_a = \mathcal{X}_\pi$). It was not known whether the space $\mathcal{X}_\lambda (\mathcal{X}_\pi, \mathcal{X}_\pi)$ might be supposed to be metrizable. The positive answer to this question is given by the following

THEOREM. *There exists a set Z of irrational numbers such that, for every formula a from the Lewis functional calculus $\mathcal{S}_\lambda^*$ (from the Heyting functional calculus $\mathcal{S}_\pi^*$, from the positive functional calculus $\mathcal{S}_\pi^*$), a is provable in $\mathcal{S}_\lambda^*$ (in $\mathcal{S}_\pi^*$, in $\mathcal{S}_\pi^*$) if, and only if, $(I_0, \mathbf{C}(Z))\Phi_a = Z$ (if $(I_0, \mathbf{H}(Z))\Phi_a = Z$).*

The set Z also fulfils the analogous theorem for the systems $\mathcal{S}_\mu^*$ and $\mathcal{S}_\pi^*$ examined in [1].

The proof of this Theorem is based on the following two topological lemmas, the second of which is a slight modification of a theorem of Albert Schwarz:

(i) Let q be a mapping of a topological space $\mathcal{X}$ into a topological space $\mathcal{Y}$. The mapping h

$$h(A) = q^{-1}(A) \quad \text{for } A \subset \mathcal{Y}$$

is a closure isomorphism of $\mathbf{C}(\mathcal{Y})$ into $\mathbf{C}(\mathcal{X})$ (i. e. a Boolean isomorphism preserving the closure operation) if, and only if, q is an interior mapping of $\mathcal{X}$ onto $\mathcal{Y}$ ([2] p. 14).

(ii) If $\mathcal{X}$ is a topological space (in the sense of [3], p. 37) with an enumerable open basis, $X \leq 2^{\aleph_0}$, then $\mathcal{X}$ is an interior image of a set of irrational numbers.

The complete proof will be published in *Fundamenta Mathematicae*.

INSTITUTE OF MATHEMATICS, POLISH ACADEMY OF SCIENCES

REFERENCES

- [1] H. Rasiowa and R. Sikorski, *Algebraic treatment of the notion of satisfiability*, *Fund. Math.* **40** (1953), 62.
- [2] R. Sikorski, *Closure homomorphisms and interior mappings*, *Fund. Math.* **41** (1954), 12.
- [3] P. Alexandroff und H. Hopf, *Topologie*, Berlin 1935.

On Some Tauberian Theorems

by

Z. POLNIAKOWSKI

Presented by W. ORLICZ on July 23, 1956

In this note we shall give the generalisations of the theorems about the Euler difference and differential equations, formulated in a previous paper [1], and some Tauberian theorems as applications.

THEOREM A. *We consider the difference equation*

$$(*) \quad L(x_n) = L^*(s_n),$$

(see [1] p. 245). If $R(r_v) \neq R(a)$ for $v=1, 2, \dots, k$, then under the hypothesis a) $\lim |s_n n^{-a}| \leq M$, or b) $s_n \sim s n^a$, the difference equation (*) has a solution $\{x_n\}$ such that a) $\overline{\lim} |\bar{x}_n n^{-a}| \leq MK$, (where K does not depend on $\{s_n\}$) or b) $\bar{x}_n \sim s \frac{W_1(a)}{W(a)} n^a$. If $W(a) \neq 0$, then the condition $R(r_v) < R(a)$, $v=1, 2, \dots, k$, is necessary and sufficient that $\overline{\lim} |x_n n^{-a}| \leq MK$ in the case a), or $x_n \sim s \frac{W_1(a)}{W(a)} n^a$ in the case b) for every solution $\{x_n\}$ of (*).

By the use of the above theorem we may prove the following Tauberian theorem: Suppose that

- 1) $w(z)$ is a polynomial of degree $k \geq 1$ and $w(n) \neq 0$ for $n=0, 1, 2, \dots$;
- 2) $R(a) > \max_v R(\varrho_v)$, where ϱ_v are zeros of $w(z)$;
- 3) $\overline{\lim} |n^{l-a} \Delta^l s_n| \leq M$ with a positive integer l ;
- 4) $t_n \sim s n^a$, where $\{t_n\}$ is the Hausdorff transform of the sequence $\{s_n\}$ corresponding to $\mu_n = 1/w(n)$.

Then $s_n \sim s w(a) n^a$.

It is easy to prove, using Theorem A, that under the hypothesis $W(n) \neq 0$ the Hausdorff transformation corresponding to $\mu_n = W_1(n)/W(n)$, preserves the asymptotic relation (more precisely, $s_n \sim s n^a$ implies $t_n \sim s \frac{W_1(a)}{W(a)} n^a$) if and only if $R(a) > \max_v R(r_v)$. Also the Mercerian theorem given in the previous note may be generalised for the case in which the convergence is replaced by the asymptotic relation mentioned above.

Theorem A can be deduced, after a little modification of the proof, from the following

THEOREM B. *The functional difference equation*

$$(**) \quad L(y) = L^*(f(x)), \quad x \geq a,$$

where

$$L(y) = \sum_{v=0}^k \lambda_v \binom{x}{v} A^v y(x-v), \quad L^*(y) = \sum_{v=0}^l \eta_v \binom{x}{v} A^v y(x-v), \quad \lambda_v = \Delta^v W(0),$$

$$\eta_v = \Delta^v W_1(0) \quad \text{and} \quad R(r_v) \neq R(a),$$

has for $x \geq a+k$ a solution $\bar{y}(x)$ such that $\overline{\lim} |\bar{y}(x)x^{-a}| \leq MK^*$ if $\overline{\lim} |f(x)x^{-a}| \leq M$, or such that $\bar{y}(x) \sim s \frac{W_1(a)}{W(a)} x^a$ if $f(x) \sim s x^a (x \rightarrow \infty)$.

If $W(a) \neq 0$, then the condition $R(r_v) < R(a)$, $v = 1, 2, \dots, k$, is necessary and sufficient that $\overline{\lim} |y(x)x^{-a}| \leq MK$, or $y(x) \sim s \frac{W_1(a)}{W(a)} x^a$ as $x \rightarrow \infty$ for every solution $y(x)$ of $(**)$ bounded in the interval $\langle x_0, x_0 + k \rangle$ with a sufficiently large $x_0 \geq a$.

In a similar way we may prove

THEOREM C. *The differential equation*

$$(***) \quad L_1(y) = L_1^*(f(x)), \quad x \geq a,$$

where

$$L_1(y) = \sum_{v=0}^k \lambda_v \frac{x^v}{v!} y^{(v)}, \quad L_1^*(y) = \sum_{v=0}^l \eta_v \frac{x^v}{v!} y^{(v)}, \quad \lambda_v = \Delta^v W(0),$$

$$\eta_v = \Delta^v W_1(0) \quad \text{and} \quad R(r_v) \neq R(a), \quad v = 1, 2, \dots, k,$$

has for $x \geq a$ a solution $\bar{y}(x)$ such that $\overline{\lim} |\bar{y}(x)x^{-a}| \leq MK$ if $\overline{\lim} |f(x)x^{-a}| \leq M$, or such that

$$\bar{y}(x) \sim s \frac{W_1(a)}{W(a)} x^a \quad \text{if} \quad f(x) \sim s x^a (x \rightarrow \infty).$$

If $W(a) \neq 0$, then the condition $R(r_v) < R(a)$, $v = 1, 2, \dots, k$, is necessary and sufficient that every solution $y(x)$ of $(***)$ should have one of the above properties respectively.

By the use of Theorem C the following Tauberian theorem may be proved: Suppose that 1') $f(x)$ has for $x \geq 0$ the continuous derivative of order $k+l$, 2') $R(a) > -1$, 3') $\overline{\lim} |x^{l-a} f^{(l)}(x)| \leq M$, 4') $C_k(x) \sim s x^a$ as

*) In this paper $\overline{\lim} \varphi(x)$ means $\overline{\lim}_{x \rightarrow \infty} \varphi(x)$.

$x \rightarrow \infty$, where

$$C_k(x) = \frac{k}{x^k} \int_0^x (x-t)^{k-1} f(t) dt \quad \text{for } x > 0.$$

Then $f(x) \sim s \binom{k+a}{k} x^a$ as $x \rightarrow \infty$.

INSTITUTE OF MATHEMATICS, POLISH ACADEMY OF SCIENCES

REFERENCES

- [1] Z. Polniakowski, *On certain theorems of the Mercer type*, Bull. Acad. Polon. Sci. Cl. III, **4** (1956), 243.

On Random Elements in Orlicz Spaces

by

A. T. BHARUCHA-REID

Presented by H. STEINHAUS on August 6, 1956

1. Let $(\Omega, \mathfrak{a}, p)$ be a probability space, that is, Ω is a space of elementary events ω , $\mathfrak{a}$ is a σ -algebra of subsets of Ω , and p is a countably additive measure on $\mathfrak{a}$ with $p(\Omega)=1$. Now let $x=x(\omega)$ be a function on Ω to $\mathfrak{X}$. Following Mourier [4], we call $x(\omega)$ a *random element* in $\mathfrak{X}$. Hanš [1] has called $x(\omega)$ a *generalised random variable*. In this note we begin the study of random elements or generalised random variables in Orlicz spaces, that is when $\mathfrak{X}=L_\Phi(\Omega, p)$.

2. Let $v = \varphi(u)$ and $u = \psi(v)$ be non-decreasing functions, inverse to each other, and satisfying $\varphi(0)=\psi(0)=0$. For given $\varphi(u)$ and $\psi(v)$, $u \geq 0$, $v \geq 0$, the convex functions $\Phi(u)$ and $\Psi(v)$ defined by the Lebesgue integrals

$$\Phi(u) = \int_0^u \varphi(r) dr, \quad \Psi(v) = \int_0^v \psi(r) dr$$

are complementary in the sense of Young. The *Orlicz space* of random elements $\mathfrak{X}=L_\Phi$ is defined as the collection of complex random elements $x(\omega)$, p -measurable on Ω , for which $\|x\|_\Phi = \sup \int_\Omega |xy| dp < \infty$, where $\sup$ is for all $y(\omega)$ such that $\int_\Omega \psi(|y|) dp \leq 1$. The conjugate Orlicz space L_Ψ is similarly defined as the collection of p -measurable random elements $x(\omega)$ such that $\|x\|_\Psi < \infty$ for $\int_\Omega \Phi(|y|) dp \leq 1$. The functions $\Phi(u)$ and $\Psi(v)$ play the same role as conjugate exponents in the Lebesgue spaces L_p (cf. [3], [6], [7]).

3. Let $\mathfrak{X}^*=L_\Phi^*$ be the adjoint or dual space of $\mathfrak{X}$, that is $\mathfrak{X}^*$ is the space of all bounded linear functionals $x^*(x)$ on $\mathfrak{X}$. The *expectation* of the random element $x(\omega)$ is the element $m = \mathcal{E}x(\omega) \in \mathfrak{X}$, if one exists, such that for all $x^* \in \mathfrak{X}^*$

$$x^*(m) = \mathcal{E}(x^*(x)).$$

Hence it follows that $\mathcal{E} x$ is the Pettis integral [2], that is

$$m = \mathcal{E} x = \int_{\Omega} x(\omega) dp,$$

since the function $x(\omega)$ on Ω to $\mathfrak{X}$ is Pettis integrable over Ω if, and only if, there exists an element $m \in \mathfrak{X}$ such that for all $x^* \in \mathfrak{X}^*$

$$\int_{\Omega} x^*(x(\omega)) dp = x^*(m).$$

For every $x \in \mathfrak{X}$ the functional $x^*(x)$ is of the form

$$x^*(x) = \int_{\Omega} x(\omega) y(\omega) dp,$$

where $y(\omega) \in L_p$ is a uniquely determined random element. The norm of x^* satisfies

$$\|y\|_{p/a} \leq \|x^*\| \leq \|y\|_p.$$

4. Hanš [1] and Mourier [4], [5] have studied strong laws of large numbers for random elements in Banach spaces. We now state two theorems for random elements in Orlicz spaces. The proofs, which are omitted, are based on the properties of linear functionals on Orlicz spaces. In a subsequent paper proofs of these theorems will be given. We will also consider some other limit theorems, and the ergodic properties of semi-groups of operators (Markov processes) on Orlicz spaces.

THEOREM 1. Let $\{x_i(\omega)\}_{i=1}^{\infty}$ be an infinite sequence of independent and identically distributed random elements in $\mathfrak{X} = L_{\Phi}$, and let $z_n = 1/n \sum_{i=1}^n x_i$. If (i) $\mathcal{E} x_i$ exist, say $\mathcal{E} x_i = \Phi$ (where Φ is the null element of $\mathfrak{X}$), and $\mathcal{E}\|x_i\| = k < \infty$, and (ii) the adjoint space $\mathfrak{X}^* = L^*\Phi$ is separable, then

$$\lim_{n \rightarrow \infty} z_n = \Phi$$

weakly almost surely.

THEOREM 2. Let $\{x_i(\omega)\}_{i=1}^{\infty}$ be a strictly stationary sequence of independent and identically distributed random elements in $\mathfrak{X} = L_{\Phi}$. If (i) $\mathcal{E} w_i$ exist, and $\mathcal{E}\|x\|$ is finite, (ii) the probability measure p is separable, and (iii) there exists a constant M such that $\Phi(2u)/\Phi(u) \leq M$, $u \geq 0$, and $\Psi(2v)/\Psi(v) \leq M$, $v \geq 0$, then

$$\lim_{n \rightarrow \infty} z_n = z \in \mathfrak{X}$$

weakly almost surely.

5. Remarks. a. Since L_{Φ} contains the same random elements as L_p when $\Phi(u) = au^p/p$, $a > 0$, $1 \leq p < \infty$, the above theorems yield laws of large numbers for random elements in L_p spaces as a special case. For L_p spaces the interpretation in terms of the existence of the p -th mo-

ments of the associated random elements is well-known. A similar interpretation in terms of the convex function $\Phi(u)$ is not known.

b. For fixed $\Phi(u)$ the space L_Φ of random elements can be made into a complete Banach space as follows: The random elements $x_1(\omega)$ and $x_2(\omega)$ are said to be in the same class if, and only if, $\|x_1(\omega) - x_2(\omega)\|_\Phi = 0$. Denoting by C the class of random elements such that $x(\omega) \in C$, then the set of random elements C form a Banach space with norm $\|C\| = \|x\|_\Phi$.

c. Luxemburg [3] has pointed out that it is possible to prove that L_Φ is uniformly convex if $\Phi(u)$ is strictly convex and there exists a constant $M > 0$ such that $\Phi(2u)/\Phi(u) \leq M$ for all $u \geq 0$. Convex Orlicz spaces may find application in the theory of statistical decision functions.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF OREGON, EUGENE, OREGON

REFERENCES

- [1] O. Hanš, *The strong law of large numbers for generalised random variables*, Bull. Acad. Polon. Sci., Cl. III **4** (1956), 15.
- [2] E. Hille, *Functional analysis and semi-groups*, New York, 1948.
- [3] W. A. J. Luxemburg, *Banach function spaces*, Assen, 1955.
- [4] E. Mourier, *Éléments aléatoires dans un espace de Banach*, Ann. de l'Inst. H. Poincaré **13** (1953), 161.
- [5] — *L-random elements and L*-random elements in Banach spaces*, Proc. Third Berkeley Symposium on Math. Statistics and Probability **2** (1956), (in press).
- [6] W. Orlicz, *Über eine gewisse Klasse von Räumen vom Typus B*, Bull. Acad. Polon. Sci. et Lettres, Série A (1932), 207.
- [7] A. C. Zaanen, *Linear analysis*, New York, 1953.

Sur l'impossibilité de définir la limite topologique inférieure à l'aide des opérations dénombrables de l'Algèbre de Boole et de l'opération de fermeture

par

R. ENGELKING

Présenté par K. KURATOWSKI le 10 Août 1956

La notion de la limite topologique supérieure et inférieure dans l'espace métrique se définit de la manière suivante:

Le point p appartient à la limite inférieure de la suite des ensembles $A_1, A_2, \dots$ ($p \in \text{Li } A_n$) si chaque entourage du point p possède des points communs avec chacun des ensembles A_m pour m suffisamment grand ([2], p. 241).

Le point p appartient à la limite supérieure de la suite des ensembles $A_1, A_2, \dots$ ($p \in \text{Ls } A_n$) si chaque entourage du point p possède des points communs avec une infinité des A_n ([2], p. 243).

La limite supérieure dans les espaces métriques satisfait à la formule suivante:

$$\text{Ls } A_n = \bigcap_{n=1}^{\infty} \sum_{m=1}^n A_{m+n},$$

donnée par Hausdorff ([1], p. 237, [4]).

Le but du travail présent est de démontrer que le cas de $\text{Li } A_n$ est entièrement différent: la notion de $\text{Li } A_n$ ne peut pas être définie à l'aide des opérations finies ou dénombrables de l'Algèbre de Boole et de l'opération de fermeture, ce qui du reste sera précisé plus exactement au § 2 et nous donnera la solution du problème posé par K. Kuratowski ([3], problème 18, et [2], p. 243).

1. LEMME. Soit $\mathcal{X}$ un ensemble dénombrable et soit $\mathcal{R}$ une famille dénombrable de ses sous-ensembles. Il existe alors un sous-ensemble infini E de l'ensemble $\mathcal{X}$ tel, que la différence $Z - E$ est infinie pour chaque ensemble infini $Z \in \mathcal{R}$.

Démonstration. On peut évidemment admettre que la famille R est composée de sous-ensembles infinis. Soit $\mathfrak{R} = \{A_1, A_2, A_3, \dots\}$.

A_i peut être représenté sous la forme d'une somme de deux ensembles infinis disjoints: $A_i = B_i + C_i$.

Posons $C_i = \{c_1^i, c_2^i, \dots\}$. Nous définissons par induction la suite $a_1, a_2, \dots$ en admettant que:

1° a_1 est un élément arbitraire de B_1 ;

2° a_i est un élément de B_i tel, que

$$(1) \quad a_i \neq a_j \neq c_m^j \quad \text{pour } j < i \quad \text{et} \quad m \leq 2i.$$

L'ensemble B_i étant infini, un tel élément a_i existe pour tout $i = 2, 3, \dots$

Posons $E = \{a_1, a_2, \dots\}$. Évidemment, E est un sous-ensemble infini de $\mathfrak{X}$.

Nous allons démontrer que $A_j - E$ est infini. Il suffit de montrer que $C_j - E$ est infini.

Supposons, par contre, que cet ensemble soit fini. Il existe alors un n tel, que pour tout $m > 2n$ on a $c_m^j \in E$.

Nous aurons alors, en particulier, pour les $4n$ termes successifs

$$c_{2n+1}^j, \dots, c_{6n}^j \in E.$$

Or, cela est impossible, car parmi les termes c_m^j aux indices $m \leq 6n$ c'est tout au plus $3n$ qui peuvent appartenir à E ; en effet, l'égalité $c_m^j = a$ entraîne en vertu de (1), $2i < m$, d'où $i < 3n$.

2. Admettons qu'il existe une formule pour $\text{Li } A_n$ écrite à l'aide $\sum_{n=1}^{\infty}$ des opérations suivantes: $+$, $\sum_{m=1}^{\infty}$, $'$, $-$, à partir des ensembles $A_{f_1(a_1, \dots, a_k) \dots A_{f_v(a_1, \dots, a_k)}}$, où f_j sont des fonctions données de k variables naturelles et aux valeurs naturelles (A' est l'ensemble complémentaire de A par rapport à $\mathfrak{X}$).

Cela signifie qu'il existe un système d'expressions

$$W_0, W_1, \dots, W_s$$

qui dépendent d'un certain nombre (≥ 0) de paramètres naturels et où:

(1) W_0 a la forme $A_{f(a_1, \dots, a_k)}$;

(2) W_l pour $l > 0$ est ou bien de la forme (1), ou bien il existe des indices $i, j < l$, tels que: soit $W_l = W_i + W_j$, soit $W_l = W_i'$, soit $W_l = \overline{W_i}$ ou enfin $W_l = \sum_{m=1}^{\infty} W_i$ (m étant un des paramètres dont dépend W_i);

(3) W_s est une formule pour $\text{Li } A_n$.

Remarque. Si au lieu de $\text{Li}_{n=\infty} A_n$ on envisageait $\text{Ls}_{n=\infty} A_n$, on aurait

$$f(n_1, n_2) = n_1 + n_2, \quad U_0 = A_f(n_1, n_2), \quad U_1 = \sum_{n_2} U_0, \quad U_2 = \bar{U}_1,$$

$$U_3 = U'_2, \quad U_4 = \sum_{n_1} U_3, \quad U_5 = U'_4 = \prod_{n_1} \prod_{n_2} \sum_{n=1} A_{n_1+n_2} \text{Ls } A_n.$$

Examinons deux substitutions pour $\mathfrak{X}$, A_n, W_i , etc. que nous désignerons respectivement $\mathfrak{X}^1, A_n^1, W_i^1$, etc. et $\mathfrak{X}^2, A_n^2, W_i^2$, etc.

I-re substitution:

$$A_n^1 = \{1/n\}, \quad \mathfrak{X}^1 = \sum_{n=1}^{\infty} A_n^1 + \{0\}.$$

Désignons par $\mathfrak{R}_i$ (pour $i = 1, 2, \dots, s$) la famille de tous les ensembles déterminés par la formule W_i^1 pour toutes les valeurs possibles de paramètres, c'est-à-dire

$$Z \in \mathfrak{R}_i = \sum_x \{Z = W_i^1(x)\},$$

où x représente l'ensemble des paramètres de la formule W_i .

Posons $\mathfrak{R} = \mathfrak{R}_1 + \dots + \mathfrak{R}_s$. Évidemment, les ensembles appartenant à la famille $\mathfrak{R}$ sont des sous-ensembles de l'espace $\mathfrak{X}^1$ et la famille $\mathfrak{R}$ est dénombrable.

Soit donc E un ensemble satisfaisant aux conditions du lemme.

Nous définissons pour $\mathfrak{X}^1$ la fonction g comme suit:

$$(4) \quad g\left(\frac{1}{n}\right) = \begin{cases} n & \text{si } \frac{1}{n} \in E \\ \frac{1}{n} & \text{si } \frac{1}{n} \notin E \end{cases} \quad g(0) = 0.$$

II-me substitution: $A_n^2 = g(A_n^1), \quad \mathfrak{X}^2 = g(\mathfrak{X}^1).$

Nous démontrerons que pour chaque $i = 1, 2, \dots, s$ et pour chaque système de paramètres de la formule W_i nous avons

$$(5) \quad g(W_i^1) = W_i^2,$$

dont on déduira aussitôt une contradiction, car l'ensemble $\text{Li}_{n=\infty} A_n^2 = W_s^2$ est vide, tandis que l'ensemble $\text{Li}_{n=\infty} A_n^1 = W_s^1$ ne l'est pas, étant composé d'un seul élément, notamment du nombre 0.

Procédons par induction —

1. Admettons que W_i a la forme (1). Pour des valeurs données des paramètres $n_1, \dots, n_k$, posons $n = f_j(n_1, \dots, n_k)$. On a alors $W_i^1 = A_n^1$, mais puisque $W_i^2 = A_n^2 = g(A_n^1)$, la formule (5) est satisfaite.

2. Admettons que W_i a la forme (2) et que pour les indices i et j la formule (5) est satisfaite. Comme g transforme d'une manière biuni-

voque l'espace $\mathfrak{X}^1$ en l'espace $\mathfrak{X}^2$, on a (pour les sous-ensembles de l'espace $\mathfrak{X}^1$) les formules suivantes:

$$g(X+Y)=g(X)+g(Y), \quad g(X')=[g(X)]', \quad g\left(\sum_{n=1}^{\infty} X_n\right)=\sum_{n=1}^{\infty} g(X_n),$$

d'où il résulte immédiatement qu'on a (pour des valeurs de paramètres établis)

$$g(W_i^1+W_j^1)=g(W_i^1)+g(W_j^1)=W_i^2+W_j^2,$$

$$g[(W_i^1)']=[g(W_i^1)]'=(W_i^2)'$$

$$g\left[\sum_{m=1}^{\infty} W_i^1\right]=\sum_{m=1}^{\infty} g(W_i^1)=\sum_{m=1}^{\infty} W_i^2.$$

La formule (5) est donc satisfaite pour W_i aux cas où W_i prend d'une des trois formes:

$$W_i=W_i+W_j, \quad W_i=W_i', \quad W_i=\sum_{m=1}^{\infty} W_i.$$

Il ne reste donc qu'à examiner le cas où $W_i=\overline{W}_i$.

Si, pour des valeurs données des paramètres, l'ensemble W_i^1 est fini, alors $W_i^1=W_i^1$. Donc

$$g(W_i^1)=g(W_i^1)=W_i^2=\overline{W}_i^2=W_i^2.$$

Admettons donc, que l'ensemble W_i^1 est infini.

En conséquence,

- (6) $W_i^1=W_i^1+\{0\}$ et l'ensemble W_i^1-E est infini; cela signifie qu'il existe une suite infinie $n_1 < n_2 < \dots$ telle que $\frac{1}{n_i} \in W_i^1-E$. En vertu de (4) $g\left(\frac{1}{n_i}\right)=\frac{1}{n_i}$, donc $\frac{1}{n_i} \in g(W_i^1-E) \subset g(W_i^1)$, c'est-à-dire $\frac{1}{n_i} \in W_i^2$ (en vertu (5)). Il en résulte que

$$(7) \quad \overline{W}_i^2=W_i^2+\{0\}.$$

Appliquant les formules (6), (5) et (7), nous arrivons finalement à la conclusion que

$$g(W_i^1)=g(W_i^1)+g(\{0\})=W_i^2+\{0\}=\overline{W}_i^2=W_i^2.$$

L'auteur tient à remercier le professeur K. Kuratowski pour ses remarques qui ont lui été très utiles dans l'accomplissement de ce travail.

INSTITUT MATHÉMATIQUE DE L'ACADÉMIE POLONAISE DES SCIENCES

OUVRAGES CITÉS

- [1] F. Hausdorff, *Grundzüge der Mengenlehre*, Leipzig, 1914.
 [2] K. Kuratowski, *Topologie I*, Monogr. Mat. **20**, Warszawa-Wrocław 1948.
 [3] Coll. Math. **1** (1947), p. 29, Problèmes.

On Models of Axiomatic Set-Theory

by

A. MOSTOWSKI

Presented on August 24, 1956

The aim of this note is to formulate a number of results concerning models of axiomatic set-theory *). Full proofs of these results will be published elsewhere.

1. Terminology and notation. Let (S) be an axiomatic system of set-theory described in [4]. (S) is based on the functional calculus of the first order with identity and has the binary relation ϵ as its unique primitive notion. The conjunction of the axioms of (S) will be denoted by $\Phi(\epsilon)$.

Most definitions and theorems to be given below belong to (S) or, speaking more accurately, can be translated to (S) by means of the technique of Gödel numbers. A number of results, however, is provable only in a stronger system (S^*) differing from (S) by having as an additional primitive a binary relation $\text{Stsf}(x, y)$ (to be read: the sequence y satisfies the formula with the Gödel number x) and additional axioms expressing the well-known properties of the relation Stsf (cf. [10], p. 271-273). The set of true formulas of (S) is known to be definable in (S^*) . Speaking in (S^*) of formulas of (S) the formulas will be identified with their Gödel numbers. Definitions and theorems of (S^*) will be distinguished from those of (S) by an asterisk.

A number of theorems and of definitions will have the form of schemes in which one or more letters denoting formulas of (S) will occur. The meaning of a scheme is that it becomes a definition or a theorem of (S) after the letters denoting formulas are replaced by arbitrary formulas of (S) .

Let $m, n, p, \dots$ be number variables, i. e. variables running over the set of integers (defined in (S) in the well-known way) and $\alpha, \beta, \gamma, \dots$

*) Parts of these results were contained in a paper read before the Section of Logic and the Foundations of Mathematics of the Third Congress of Soviet Mathematicians in Moscow on July 2, 1956.

function variables, i. e., variables running over the set of infinite sequences of integers. We denote by **A** the class of arithmetical formulas, i. e. the least class containing formulas of the form

$$(1) \quad m=n, \quad m=n+p, \quad m=n \cdot p,$$

and satisfying the conditions

(2) if A, B are in **A**, then so are $\sim A, A \vee B, (m)A, (\exists m)A$, where m is any number variable.

We denote further by **E** the class of elementary formulas, i. e. the least class containing the formulas (1) as well as formulas of the form $m=a(n)$ and satisfying the conditions (2). Note that function variables are never bound in elementary formulas *).

Strictly speaking, **A** and **E** are not classes of formulas of (S) since in (S) all the variables are of the same type. It is easy, however, to translate arithmetical and elementary formulas to the language of (S) , and we shall henceforth identify **A** and **E** with classes of their translations.

If $A(\epsilon)$ is an arbitrary formula of (S) (which may contain free variables), and x is a variable not occurring in $A(\epsilon)$, then we denote by $A(x)$ the formula obtained from $A(\epsilon)$ by replacing each subformula of the form $u \in v$ by $\langle u, v \rangle \in x$ and limiting all quantifiers to the set $\text{Field } (x) = E_s(\exists t)((\langle s, t \rangle \in x) \vee (\langle t, s \rangle \in x))$.

2. Models. DEFINITION 1. $\text{Mod}(R) = (R \text{ is a set of ordered pairs}) \& \Phi(R)$.

The formula $\text{Mod}(R)$ can be read: R is a model of (S) . If $\text{Mod}(R)$, then R has a unique minimal element 0_R . For arbitrary $s, t \in \text{Field}(R)$ there is exactly one element z such that $(x)(xRz \supset ((x=s) \vee (x=t)))$. We denote this z by $\{s, t\}_R$ and put $\langle s, t \rangle_R = \{\{s, t\}_R, \{t, s\}_R\}$.

DEFINITION 2. $\text{Fund}(R) = 0$ if $\sim \text{Mod}(R)$ and $\text{Fund}(R)$ is the least set X such that $0_R \in X$ and $(z)((x)(xRz \supset x \in X) \supset z \in X)$ if $\text{Mod}(R)$. $\text{Fund}(R)$ is called the well-founded part of R .

LEMMA. If $\text{Mod}(R)$, then the relation R limited to $\text{Fund}(R)$ is isomorphic with the relation $\in$ in a complete **) set $M(R)$ of sets. There is exactly one function which establishes this isomorphism.

Proof of this lemma is given in [4], p. 147.

For $x \in \text{Fund}(R)$ we shall denote by x' the isomorphic image of x in $M(R)$.

DEFINITION SCHEMA 3. A model R is absolute for a formula $Z(\epsilon, u_1, u_2, \dots)$ if for every $u_1, u_2, \dots$ in $\text{Field}(R)$

$$Z(R, u_1, u_2, \dots) \equiv (u_1, u_2, \dots \in \text{Fund}(R)) \& Z(\epsilon, u'_1, u'_2, \dots).$$

*) Instead of elementary formulas we could have used the non-elementary arithmetical formulas of [4].

**) A set x is complete if $u \in v \in x \supset u \in x$.

Let $\text{Fin}(\epsilon, x)$ be the formula " x is a finite set" and $\text{Bord}(\epsilon, x)$ the formula " x is a well-ordering relation".

THEOREM 1 *). R is absolute for the formula $\text{Fin}(\epsilon, x)$ (abbreviated $\text{Mod}_I(R)$) if, and only if, $\text{Fin}(R, x) \supset \{ \text{the set } E_u(uRx) \text{ is finite} \}$.

THEOREM 2 *). The following three conditions are equivalent: 1° R is absolute for the formula $\text{Bord}(\epsilon, x)$; 2° $\text{Fund}(R) \equiv \text{Field}(R)$; 3° for each x in $\text{Field}(R)$ if $\text{Bord}(R, x)$, then the relation $E_{u,v}[(u, v)_R R x]$ is a well-ordering relation.

Each of the three equivalent conditions 1°-3° will be denoted by $\text{Mod}_{II}(R)$.

Instead of $\text{Mod}_I(R)$ one could also say that R is a model absolute for integers and, instead of $\text{Mod}_{II}(R)$, that R is a model absolute for ordinals. It can in fact be shown that $\text{Mod}_I(R)$ if, and only if, the set of integers of the model R is isomorphic to the set of integers and $\text{Mod}_{II}(R)$ if, and only if, each ordinal of the model R is isomorphic to an ordinal.

THEOREM 3. $\text{Mod}_{II}(R) \supset \text{Mod}_I(R)$.

In the next three theorems $Z(\epsilon)$ is a formula without free variables.

THEOREM 4. R is absolute for $Z(\epsilon)$ if, and only if, $Z(\epsilon) \equiv Z(R)$.

*THEOREM 5. The formula "for each R - if $\text{Mod}(R)$, then R is absolute for $Z(\epsilon)$ " is true if, and only if, $Z(\epsilon)$ is decidable in (S) .

*THEOREM 6. The formula "for each R - if $\text{Mod}_I(R)$, then R is absolute for $Z(\epsilon)$ " is true if, and only if, $Z(\epsilon)$ is decidable by means of the rule ω^{**} .

Theorem 6 is a corollary to a more general theorem due to Orey [5].

3. Provability of formulas expressing the existence of absolute models.

In the theorems given in this section we are investigating the problem as to which new axioms must be added to (S) , in order that the existence of absolute models should be provable. We denote by $(S) \& A$ the system obtained from (S) by the adjunction of a new axiom A .

THEOREM 7. If $Z(\epsilon)$ has no free variables and $\sim Z(\epsilon)$ is not provable in (S) , then the existence of a model absolute for $Z(\epsilon)$ is not provable in $(S) \& Z(\epsilon)$.

Thus the existence of a model absolute for all formulas is provable neither in (S) nor in any finite extension of (S) .

Rosser and Wang [8], p. 123, have proved that if (S) is consistent, then the formula $(\exists R) \text{Mod}(R) \supset (\exists R) \text{Mod}_I(R)$ is not provable in (S) . Ryll-Nardzewski generalised their result by showing the unprovability

*) Models absolute for the formulas Fin and Bord were considered in [8], pp. 113 and 123.

***) This rule is called the "rule of infinite induction" in [10], p. 259 and the "Carnap rule" in [7], p. 129.

of formulas $A \supset (\exists R) \text{Mod}_I(R)$, where A is an arithmetical formula consistent with (S) .

THEOREM 8. *There exist elementary formulas $F_I(\epsilon, \alpha)$ and $F_{II}(\epsilon, \alpha, \beta)$ such that $(\exists R) \text{Mod}_I(R) = (\exists \alpha) F_I(\epsilon, \alpha)$ and $(\exists R) \text{Mod}_{II}(R) = (\exists \alpha)(\beta) F_{II}(\epsilon, \alpha, \beta)$.*

The construction of F_{II} is given in [4], p. 154-156.

THEOREM SCHEMA 9. *For every elementary formula $G(\epsilon, \gamma_1, \gamma_2, \dots, \gamma_k)$, if $\text{Mod}_{II}(R)$, then R is absolute for the formula $(\exists \gamma_1, \dots, \gamma_k) G(\epsilon, \gamma_1, \dots, \gamma_k)$.*

The proof of this theorem can be based on the results either of Kleene [2], or of Sierpiński [9].

COROLLARY 10. *If $G_i(\epsilon, \gamma_1, \dots, \gamma_k)$ and $H_i(\epsilon, \gamma_1, \dots, \gamma_k)$ ($i=1, \dots, n$) are elementary formulas, then the formula $(\exists R) \text{Mod}_{II}(R)$ is provable in the system*

$$(S) \& \prod_{i=1}^n ((\exists \gamma_1, \dots, \gamma_k) G_i(\epsilon, \gamma_1, \dots, \gamma_k) \& \prod_{i=1}^n (\gamma_1, \dots, \gamma_k) H_i(\epsilon, \gamma_1, \dots, \gamma_k) *),$$

if, and only if, this system is inconsistent. In particular the formula $(\exists R) \text{Mod}_I(R) \supset (\exists R) \text{Mod}_{II}(R)$ is not provable in (S) if the system $(S) \& (\exists R) \text{Mod}_I(R)$ is consistent.

THEOREM SCHEMA 11. *For every elementary formula $G(\epsilon, \gamma_1, \dots, \gamma_k)$ there exists an ordinal σ_G such that if 1° $\text{Mod}_I(R)$ and 2° $\text{Field}(R)$ contains an element x with the property that the relation $E_{u,v}[(u, v)_R R x]$ is of type σ_G , then R is absolute for the formula $(\exists \gamma_1, \dots, \gamma_k) G(\epsilon, \gamma_1, \dots, \gamma_k)$.*

The proof of this theorem is based on Sierpiński [9].

THEOREM SCHEMA 12. $\sigma_G \leq \omega_1$, where ω_1 is the first non-constructive ordinal.

The proof follows the lines of Kleene's [2], pp. 415-419.

***THEOREM 13.** *If the formula $(\exists \gamma_1, \dots, \gamma_k) G(\epsilon, \gamma_1, \dots, \gamma_k)$ (where G is an elementary formula) is false, then $\sigma_G < \omega_1$.*

THEOREM SCHEMA 14. *If σ_G is less than ω_1 , then the formula $(\exists \gamma_1, \dots, \gamma_k) G(\epsilon, \gamma_1, \dots, \gamma_k)$ is decidable by means of the rule ω . * In particular, each true formula of the form $(\gamma_1, \dots, \gamma_k) G(\epsilon, \gamma_1, \dots, \gamma_k)$, where G is an elementary formula, is provable by means of this rule.*

THEOREM 15. *There are elementary formulas $G(\epsilon, \alpha)$ with $\sigma_G = \omega_1$.*

An example of such a formula is furnished by Rosser [7], Theorem 12.

***THEOREM 16.** *There is an elementary formula $A(\epsilon, \alpha, \beta, \gamma)$ such that the following formula is false: "there is an ordinal σ such that if 1° $\text{Mod}_{II}(R)$ and 2° $\text{Field}(R)$ contains an x for which the relation $E_{u,v}[(u, v)_R R x]$ is of order type σ , then R is absolute for the formula $(\alpha)(\exists \beta)(\gamma) A(\epsilon, \alpha, \beta, \gamma)$ ".*

*) II stands for the symbol of conjunction.

It follows from Theorem 16 that Theorem 11 cannot be generalised to formulas beginning with the quantifiers $(\alpha)(\exists\beta)(\gamma)$.

4. Pathological models *. Let L be the Lindenbaum algebra (cf. [6], p. 197) of the system (S) and L_∞ the Lindenbaum algebra of the system (S_∞) obtained from (S) by closing it with respect to the rule ω . Let S and S_∞ be the Stone spaces of algebras L and L_∞ . Each element I of S or of S_∞ defines a pseudo-model of (S) or of (S_∞) . The sets of those I which define models of these systems are according to [6], p. 197 of second category in the respective spaces and hence non-void.

THEOREM 17. *The set of I in S , which define models R such that R limited to $\text{Fund}(R)$ is not a model of (S) , is of second category in S .*

COROLLARY 18. *There are denumerable models R of (S) such that R limited to $\text{Fund}(R)$ is not a model of (S) .*

***THEOREM 19.** *The set of I in S_∞ which define well-founded models of (S_∞) , is nowhere dense in S_∞ .*

From Theorem 19 and from the proof of the theorem of Orey referred to on p. 667 we also obtain the following

***COROLLARY 20.** *There are models R such that $\text{Mod}_I(R) \& \sim \text{Mod}_{II}(R)$.*

REFERENCES

- [1] S. C. Kleene, Trans. Amer. Math. Soc. **79** (1955), 312.
- [2] — Am. Journal of Math. **77** (1955), 405.
- [3] A. Mostowski, Studia Philosophica **4** (1951), 237.
- [4] — Fund. Math. **36** (1949), 143.
- [5] S. Orey, Journ. Symb. Logic **21** (1956), 246.
- [6] H. Rasiowa and R. Sikorski, Fund. Math. **37** (1950), 193.
- [7] J. B. Rosser, Journ. Symb. Logic **2** (1937), 129.
- [8] J. B. Rosser and Hao Wang, Journ. Symb. Logic **15** (1950), 113.
- [9] W. Sierpiński, Fund. Math. **8** (1926), 362.
- [10] A. Tarski, *Logic, semantics, metamathematics*, Oxford, 1956, p. 152.

*) Proofs of theorems given in this section make use of topological ideas of [6]. That these ideas can be applied to this type of problems has been suggested by Ryll-Nardzewski in the proof given by him of Orey's theorem mentioned on p. 667.

Character of Interaction Potential Between Electron and Photon

by

J. WRZECIONKO

Presented by W. RUBINOWICZ on July 10, 1956

Electron-photon scattering is discussed on the basis of the Tamm-Dancoff method. The electron-photon interaction potential is found in the Dyson approximation [2]. The angular variables and the various spin states are separated, and the interaction potential is investigated for the various spin states. An analogy with the nucleon-pion potential [2] is established.

In the system of units used throughout this paper $\hbar=1=c$.

Notation:

$a_r^+(\vec{p})$, $a_r(\vec{p})$ — the respective creation and annihilation operators for an electron of momentum $\vec{p}$ and spin r .

$b_r^+(\vec{p})$, $b_r(\vec{p})$ — the analogous operators for a positron,

$C_\lambda^+(\vec{k})$, $C_\lambda(\vec{k})$ — creation and annihilation operators for a photon of polarization λ and momentum $\vec{k}$.

The basis of the problem is provided by the standing wave equation [1]

$$(1) \quad \Psi_a^{(1)} = \Psi_a + P \left[\frac{1}{E - H_0} \right] H_1 \Psi_a^{(1)},$$

where H_1 and H_0 are the interaction and free field Hamiltonians in the Schrödinger picture and Ψ_a is one particular eigenstate of H_0 .

Equation [1] contains the initial conditions (index a).

The T. D. amplitudes are determined from the equation

$$\begin{aligned} & {}^a q_{\dots n m l}^{\dots r_i \dots r'_k \dots \lambda_j} (\dots \vec{p}_i \dots; \dots \vec{p}'_k \dots; \dots \vec{k}_j \dots) = \\ & = \left(\Psi_0, \prod_{i=1}^n a_{r_i}(\vec{p}_i) \prod_{k=1}^m b_{r'_k}(\vec{p}'_k) \prod_{j=1}^l C_{\lambda_j}(\vec{k}_j) \Psi_a^{(1)} \right), \end{aligned}$$

where Ψ_0 is the vacuum state of the free fields.

Henceforth we shall omit the index a in ${}^a q$. We obtain from [1] a set of equations for the amplitude q_{nml} . In what follows we shall be interested in the amplitude $q_{101}(\vec{p}r, \vec{k}t)$. We shall find the integral equation for this amplitude in the Dyson approximation. The Dyson approximation [2] is based on the assumption that the following amplitudes are different from zero: q_{100} , q_{210} , q_{102} , q_{212} and q_{101} . After eliminating all amplitudes except $q_{101}(\vec{p}r, -\vec{p}t) = q_{r1}(\vec{p})$, we obtain the following equations (we pass over to the system in which $\vec{p} = -\vec{k}$, $\vec{p}$ being the momentum of the electron and $\vec{k}$ — the momentum of the photon):

$$\begin{aligned}
 (E - E(\vec{p}) - \omega(\vec{p}))q_{r1}(\vec{p}) &= -\frac{e^2}{(2\pi)^{12}} P \int d_3 s \sum_{r_0 l} (4\omega(\vec{p})\omega(\vec{s})^{-\frac{1}{2}} \times \\
 &\times \bar{u}^r(\vec{p}) \left\{ \hat{e}_l^2 \frac{\lambda + (0)}{E - m} \hat{e}_l^1 + \hat{e}_l^2 \frac{\lambda - (-0)}{E - E(\vec{p}) - E(\vec{s})} \frac{\omega(\vec{p}) - \omega(\vec{s}) - m}{\omega(\vec{p}) - \omega(\vec{s}) - m} \hat{e}_l^1 + \right. \\
 (2) \quad &+ \hat{e}_l^1 \frac{\lambda + (\vec{p} + \vec{s})}{E - E(\vec{p} + \vec{s})} \frac{\omega(\vec{p}) - \omega(\vec{s})}{\omega(\vec{p}) - \omega(\vec{s})} \hat{e}_l^2 + \hat{e}_l^1 \frac{\lambda - (-\vec{p} - \vec{s})}{E - E(\vec{p} + \vec{s}) - E(\vec{p}) - E(\vec{s})} \hat{e}_l^2 \Big\} \times \\
 &\times U_{r_0}(\vec{s}) q_{r_0 l}(\vec{s}) \\
 \lambda + (\vec{p}) &= -\frac{1}{2E(\vec{p})} [i\hat{p} - m]; \quad \hat{p} = \vec{\gamma}\vec{p} + p_4 \gamma^4 = p_\mu \gamma^\mu; \quad p_4 = iE(\vec{p}) \\
 \lambda - (\vec{p}) &= \frac{1}{2E(\vec{p})} [-i\hat{p} - m]; \quad E(\vec{p}) = (m^2 + \vec{p}^2)^{\frac{1}{2}}; \quad \omega(\vec{p}) = (\vec{p}).
 \end{aligned}$$

The kernel of the integral equation (2) consists of 4 terms. The individual terms of the kernel represent the processes:

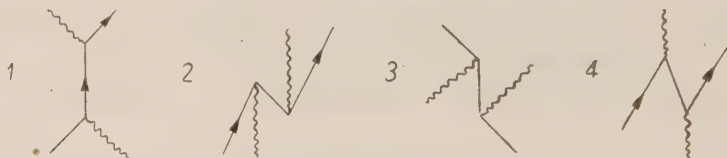


Fig. 1

respectively.

In the integral equation for q_{101} we have discarded the kernels corresponding to the self-energy of the electron, photon and vacuum $\rightarrow$ vacuum transitions, which we obtained by successive elimination of the amplitudes q_{100} ; q_{210} ; q_{102} ; q_{212} .

It will further be convenient to use the amplitude defined as follows:

$$q_{rl}(\vec{p}) = \sum_{l=1}^4 \sum_{r=1}^2 e_{l\mu}^2 U^r(\vec{p}) q_{r1}(\vec{p}),$$

where μ is vector index,
 t is the photon polarization index,
 r is the electron spin index.

Then, from equation [2] we obtain

$$\begin{aligned} (E - E(\vec{p}) - \omega(\vec{p})) \gamma_\mu(\vec{p}) = & -\frac{e^2}{(2\pi)^{12}} \frac{-\beta\gamma_5(\vec{\sigma}\vec{p}) + \beta E(\vec{p}) + m}{2E(\vec{p})} \times \\ & P \int d_3\vec{s} (4\omega(\vec{p})(\vec{s}))^{-\frac{1}{2}} \left\{ \gamma^\mu \frac{\beta+1}{2} \gamma^\nu \frac{1}{E-m} + \right. \\ & + \gamma^\mu \frac{\beta-1}{2} \gamma^\nu \frac{1}{E - E(\vec{p}) - E(\vec{s}) - \omega(\vec{p}) - \omega(\vec{s}) - m} + \\ & + \gamma^\nu \frac{-\beta\gamma_5\vec{\sigma}(\vec{p} + \vec{s}) + \beta E(\vec{p} + \vec{s}) + m}{2E(\vec{p} + \vec{s})} \gamma^\mu \frac{1}{E - E(\vec{p} + \vec{s}) - \omega(\vec{p}) - \omega(\vec{s})} + \\ & \left. + \gamma^\nu \frac{\beta\gamma_5\vec{\sigma}(\vec{p} + \vec{s}) + \beta E(\vec{p} + \vec{s}) - m}{2E(\vec{p} + \vec{s})} \gamma^\mu \frac{1}{E - E(\vec{p} + \vec{s}) - E(\vec{s}) - E(\vec{p})} \right\} q_\nu(\vec{s}). \end{aligned}$$

For each q_μ , the amplitude μ is a bispinor, and we can write it in the form

$$(4) \quad q_\mu(\vec{p}) = \begin{pmatrix} q_\mu^+(\vec{p}) \\ q_\mu^-(\vec{p}) \end{pmatrix} = \begin{pmatrix} q_\mu^+(\vec{p}) \\ 0 \end{pmatrix} + \gamma_5 \begin{pmatrix} q_\mu^-(\vec{p}) \\ 0 \end{pmatrix};$$

q^+ and q^- — being the large and small Pauli components.

We shall consider the terms $(\vec{\sigma}_1\vec{p}_1)\dots(\vec{\sigma}_n\vec{p}_n)q^\pm$. When $q^\pm$ has the $+$ sign and n is even, we shall treat it as a large Pauli component; when n is odd — as a small component. When $q^\pm$ has the $-$ sign, the situation is the reverse. It turns out that equation [3] can be split unambiguously into two equations corresponding to large q^+ and small q^- Pauli components. By suitably denoting the electromagnetic potential, we can take the fourth (time) component of the 4-vector to be equal to zero ($q_4=0$). In this case, equation [3] separates into large and small Pauli components. Henceforth, we shall make the approximation

$$(3) \quad q_\mu^-(\vec{p}) = -\frac{(\vec{\sigma}\vec{p})}{E(\vec{p}) + m} q_\mu^+(\vec{p}).$$

Taking the above remarks into account, we obtain the equation for $q_i^+(\vec{p})$; $i=1,2,3$.

Carrying out a separation of the angular variables in the equation for q_i^+ , we obtain:

$$\begin{aligned} (5) \quad & \left(E - E(p) - \omega(p) \right) q_i^+(p) = -\frac{e^2}{2^{10}\pi^{11}} P \int_0^\infty s^2 ds \left(4\omega(p)s\omega(s) \right)^{-\frac{1}{2}} \times \\ & \left\{ \sum_{j=1}^3 -\sigma_i\sigma_j \left[L_2 \frac{E(p)+m}{2E(p)} + X_{i\pm 1}^5 \frac{p \cdot s}{2l+1} + \right. \right. \\ & \left. \left. + \sigma_j\sigma_i \frac{1}{2l+1} \left[X_l^1 + X_{l\pm 1}^2 p \cdot s + X_{l\pm 1}^3 p \cdot s + X_l^3 p^2 - X_{l\pm 1}^4 p \cdot s + X_l^4 s^2 \right] \right] q_j^+(s), \right\} \end{aligned}$$

where

$$\begin{aligned}
 p &= |\vec{p}|; L_2 = \frac{1}{E - E(p) - E(s) - \omega(p) - \omega(s) - m} \\
 X_n^1 &= \left(n + \frac{1}{2}\right) \frac{1}{4E(p)} \frac{1}{r} \int_{1-r}^{1+r} z \left(1 - \frac{m}{z\bar{E}}\right) \left(\frac{E(p) + m}{z\bar{E} - B} + \frac{E(p) - m}{z\bar{E} - C}\right) P_n\left(\frac{z^2 - 1 - r^2}{2r}\right) dz, \\
 X_n^2 &= \left(n + \frac{1}{2}\right) \frac{1}{4E(p)(E(s) + m)} \frac{1}{r} \int_{1-r}^{1+r} z \left[\frac{1}{z\bar{E} - B} \left(\frac{m}{z\bar{E}} + 1\right) - \right. \\
 &\quad \left. - \frac{1}{z\bar{E} - C} \left(\frac{m}{z\bar{E}} - 1\right)\right] P_n\left(\frac{z^2 - 1 - r^2}{2r}\right) dz, \\
 (6) \quad X_n^3 &= \left(n + \frac{1}{2}\right) \frac{1}{4E(p)\bar{E}} \frac{1}{r} \int_{1-r}^{1+r} \left[\frac{1}{z\bar{E} - \beta} - \frac{1}{z\bar{E} - C}\right] P_n\left(\frac{z^2 - 1 - r^2}{2r}\right) dz, \\
 X_n^4 &= \frac{E(p) + m}{E(s) + m} X_n^3, \\
 X_n^5 &= \left(n + \frac{1}{2}\right) \frac{1}{4\bar{E}(p)(E(s) + m)(E - m)} \frac{1}{r} \int_{1-r}^{1+r} P_n\left(\frac{z^2 - 1 - r^2}{2r}\right) dz,
 \end{aligned}$$

P_n is the Legendre polynomial,

$$B = E - \omega(s) - \omega(p); \quad C = E - E(s) - E(p),$$

$$r = \frac{p \cdot s}{\bar{E}^2}; \quad \bar{E} = \frac{E(p + s) + E(p - s)}{2}.$$

We may write the integral equation in the form

$$\begin{aligned}
 (E - E(p) - \omega(s)) \vec{q}^+(p) &= \\
 (7) \quad &= e^2 P \int_0^\infty s^2 ds (4\omega(p)\omega(s))^{-1} [N_1 \hat{L}_1(p_1 s) + \\
 &\quad + \hat{N}_2 L_2(p_1 s)] X \vec{q}^+(s), \\
 \vec{q}^+(p) &= [q_1^+(p), q_2^+(p), q_3^+(p)].
 \end{aligned}$$

The matrix representation of the operators $\hat{N}_1$ and $\hat{N}_2$ are

$$(\hat{N}_1^2)_{ij} = \sigma_i \sigma_j; \quad (\hat{N}_2)_{ij} = \sigma_j \sigma_i.$$

We find that the eigenvalues of the operators $\hat{N}_1$ and $\hat{N}_2$ are

$$(8) \quad (\hat{N}_1^2)_{ij} = \sigma_i \sigma_k \sigma_k \sigma_j = 3 \sigma_i \sigma_j = 3 (\hat{N}_1)_{ij}.$$

Thus, the eigenvalues of $\hat{N}_1$ are 0 and 3.

We find the eigenvalues of $\hat{N}_2$ by employing the commutation rule for the matrix

$$(9) \quad (\hat{N}_2)_{ij} = \sigma_j \sigma_i = 2\delta_{ij} - \sigma_i \sigma_j = 2\delta_{ij} - (\hat{N}_1)_{ij}.$$

Thus, the eigenvalues of $\hat{N}_2$ are 2 and -1.

If the eigenvalue of $\hat{N}_1$ is 0, then the eigenvalue of $\hat{N}_2$ is 2.

If the eigenvalue of $\hat{N}_1$ is 3, then $\hat{N}_2$ has the eigenvalue -1 .

We shall now find the eigenvalues of the operators under consideration for the various spin states.

The kernel $L_1(p_1s)$ corresponds to processes 1 and 2. Let us consider process 1. The virtual state is the state of the electron. Therefore its spin is $1/2$. As a result, an electron of spin $\pm 1/2$ can absorb a photon with a spin of $\mp$ only.

An electron of spin $1/2$ cannot absorb a photon of spin $+1$. This state would be a state with spin $3/2$. Since $\hat{N}_1$ describes the spin states for processes 1 and 2, we ascribe the eigenvalue 0 to the spin state $3/2$, and the eigenvalue 3 to the spin state $1/2$. We thus have, single-valuedly: the eigenvalue 2 of operator $\hat{N}_2$ corresponds to the spin state $3/2$, and -1 , to the spin state $1/2$.

The integral equation can be written more briefly:

$$(10) \quad (E - E(p) - \omega(p))\vec{q}^+(p) = e^2 P \int_0^\infty s^2 ds (4\omega(p)\omega(s))^{-\frac{1}{2}} L(p_1s)\vec{q}^+(s)$$

$$L(p_1s) = \begin{cases} 3L_1(p_1s) - L_2(p_1s) & \text{for spin state } j = \frac{1}{2} \\ 2L_2(p_1s) & \text{for spin state } j = \frac{3}{2} \end{cases}$$

The kernel of the integrated equation gives a divergence of degree $5/2$. Hence the main contribution to the integral is given by the domain of integration for which $s \gg m$. This means that the electron-photon interaction potential is concentrated at very small distances. Therefore, in investigating the character of the potential, it suffices to investigate its sign for $s \gg m$.

Since our equation is taken in the Pauli approximation, it gives a true description of the processes for non-relativistic energies. Thus, as regards the initial and final momenta of the electron and photon, we make the assumption that

$$p, p_0, k_0 \ll m.$$

The solution of equation (10) can be represented in the form

$$(11) \quad \vec{\Psi}(p) = \vec{\Psi}_B(p) - e^2 P \int_0^\infty s^2 ds (4\omega(p)\omega(s))^{-\frac{1}{2}} \frac{L(p_1s)}{E(s) + \omega(s) - E} \vec{\Psi}(s)$$

$$\vec{\psi}(p) = e^2 P \int_0^\infty s^2 ds (4\omega(p)\omega(s))^{-\frac{1}{2}} L(p_1s)\vec{q}^+(s)$$

$$\vec{\psi}_B(p) = e^2 P \int_0^\infty s^2 ds (4\omega(p)\omega(s))^{-\frac{1}{2}} L(p_1s)\delta(E - E(s) - \omega(s)).$$

If $L(p, s) > 0$, then $|\vec{\Psi}| < |\vec{\Psi}_B|$; the potential will therefore be repulsive.

For $L(p,s) < 0$, $|\vec{\Psi}| > |\vec{\Psi}_B|$; the potential will be attractive. The sign of the potential $L(p,s)$ should be fixed. We can do this by means of the assumptions regarding the momenta. For this purpose we found the asymptotic behaviour of the coefficients (6):

$$(12a) \quad X_0^1 \sim \frac{1}{4s}; \quad X_0^2 \sim \frac{1}{4s^2m}; \quad X_0^3, X_0^4 \sim 0; \quad X_0^5 \sim \frac{1}{4m(E-m)s}$$

$$(12b) \quad X_1^1 \sim -\frac{1}{8} \frac{m}{sp}; \quad X_1^2, X_1^3, X_1^4 \sim 0; \quad X_1^5 \equiv 0$$

$$(12c) \quad X_2^1 \sim \frac{5S}{8pm} \left(\frac{3}{2} \frac{m}{p} + \frac{3}{8} \frac{p}{m} \right) + \frac{15}{8} \frac{s}{p^2}; \quad X_2^2 \sim \frac{12.34}{mp^2}; \quad X_2^3, X_2^4 \sim 0$$

$$(12d) \quad X_3^1 \sim -\frac{1.65}{p}; \quad X_3^2 \sim -\frac{1.65}{msp}; \quad X_3^3, X_3^4 \sim 0; \quad X_3^5 \equiv 0.$$

Using the results obtained, we investigated the sign of the kernel in different spin states

The choice of the + or - sign before the 1 in $X_{l\pm 1}^p$ is conditioned by the value of the number $j = l \pm \frac{1}{2}$.

If $j = l + \frac{1}{2}$, then X_{l-1}^p .

If $j = l - \frac{1}{2}$, then X_{l+1}^p .

I. Spin state $j = \frac{1}{2}$.

Two states are possible: for $l=0$ and $l=1$. The potential in this case is $L(p,s) = 3L_1(p,s) - L_2(p,s)$. We found that $L^{S_{1/2}}(p,s) \sim -\frac{5}{4} \frac{1}{S} < 0$, and $L^{P_{1/2}}(p,s) \sim \frac{12.34}{3} \frac{s}{pm} > 0$. Therefore, in the $S_{1/2}$ state, the potential is attractive; in the $P_{1/2}$ state, the potential is repulsive.

II. Spin state $j = \frac{3}{2}$.

States are possible: for $l=1$ and $l=2$. We found that $L^{P_{3/2}}(p,s) \sim \frac{2}{S} \left(\frac{1}{24} \frac{m}{p} - \frac{1}{12} \frac{p}{m} \right) > 0$, and $L^{D_{3/2}}(p,s) \sim -\frac{2}{5} \left[\frac{45}{16} \frac{p}{p^2} - \frac{1.65}{m} \right] < 0$.

In the $P_{3/2}$ state, the potential is repulsive.

In the $D_{3/2}$ state, the potential is attractive.

In [2], Dyson obtained partially analogous results for nucleon-pion scattering. He obtained an attractive potential for the state $P_{3/2}$ with an isotopic spin of $3/2$.

The author wishes to express his appreciation and thanks to W. Królikowski under whose direction this paper was written.

INSTITUTE OF PHYSICS, POLISH ACADEMY OF SCIENCES

REFERENCES

- [1] B. Lippman and J. Schwinger, Phys. Rev. **79** (1950), 469.
- [2] F. J. Dyson et al., Phys. Rev. **95** (1954), 1644.

On the Conservation Theorems and Equations of Motion in Covariant Field Theories

by

A. TRAUTMAN

Presented by L. INFELD on August 14, 1956

1. In this note we shall deal with a tensor field (or set of tensor fields) the components of which will be denoted by $\psi_A(x^\nu)$ ($A=1, \dots, N$; $\nu=0, \dots, 3$). The field equations for the pure field case are supposed to be derivable from a variational principle

$$(1) \quad \delta I = 0,$$

where:

$$(2) \quad I = \int_{\Omega} \mathfrak{L} d_4 x, \quad \mathfrak{L} = \sqrt{-g} L,$$

$$(3)_1 \quad L = L(g_{\mu\nu}, g_{\mu\nu,\sigma}, \psi_A, \psi_{A,\nu}), \quad g = \det(g_{\mu\nu});$$

$g_{\mu\nu}(x^\alpha)$ is the metric tensor of the space-time V_4 (in general non-Euclidean). We assume that L is an invariant with respect to general co-ordinate transformations in V_4 .

The purpose of this paper is to obtain a conservation law for the ψ -field and to derive the equations of motion for a "cluster of particles" moving in such a field. The latter problem is closely related to a paper by Infeld [1] on the equations of motion in linear field theories.

2. We use the notation of the general relativity theory (summation convention, ordinary (covariant) differentiation denoted by a comma (semicolon), $D\varphi^A \equiv \varphi^A_{;\nu} dx^\nu$).

The field equations resulting from (1) are:

$$(4) \quad L^A \stackrel{\text{def.}}{=} \frac{1}{\sqrt{-g}} \left(\frac{\partial \mathfrak{L}}{\partial \psi_A} - \frac{\partial}{\partial x^\nu} \frac{\partial \mathfrak{L}}{\partial \psi_{A,\nu}} \right) = 0.$$

We generalise (4), assuming that the sources of the field are described by a tensor field j^A ; the equations in the presence of sources become $L^A = j^A$.

We define a symmetric energy-momentum density tensor:

$$(5) \quad T_{\alpha\beta} \stackrel{\text{def.}}{=} \frac{2}{\sqrt{-g}} \left(\frac{\partial \mathcal{L}}{\partial g^{\alpha\beta}} - \frac{\partial}{\partial x^\nu} \frac{\partial \mathcal{L}}{\partial g^{\alpha\beta}{}_{,\nu}} \right).$$

It is well known [2], [3] that the conservation laws for physical systems are intimately connected with the invariance properties of the Lagrangian function. The invariance under the group of general co-ordinate transformations in V_4 leads to four *identities* connecting $T_{\alpha\beta}$, L^A and the ψ 's.

Let us take an infinitesimal co-ordinate transformation $x^\nu \rightarrow x'^\nu = x^\nu + \delta\zeta^\nu$, where $\delta\zeta^\nu(x^\alpha)$ is a vector field. We form the "substantial" variation of ψ_A :

$$(6) \quad \delta^* \psi_A \stackrel{\text{def.}}{=} \psi_A(x') - \psi_A(x^\nu), \quad \text{where} \quad x'^\nu = x^\nu + \delta\zeta^\nu = x^\nu, *$$

and assume, following Belinfante [5] and Bergmann [6], that

$$(7) \quad \delta^* \psi_A = -\psi_{A,\nu} \delta\zeta^\nu + F_{A\mu}{}^{B\nu} \psi_B \delta\zeta^\mu{}_{,\nu}.$$

$F_{A\mu}{}^{B\nu}$ is a constant (numerical) tensor depending on the transformation properties of ψ_A and fulfilling certain commutation relations [6]. From (7), we obtain:

$$(8) \quad \psi_{A;a} = \psi_{A,a} + F_{A\mu}{}^{B\nu} I_{\nu}^\mu \psi_B.$$

The original region of integration Ω is mapped, by the transformation $x^\nu \rightarrow x'^\nu$, to a new region Ω' . We can now write:

$$(9) \quad \delta^* I \stackrel{\text{def.}}{=} \int_{\Omega'} \mathcal{L}' d_4 x' - \int_{\Omega} \mathcal{L} d_4 x = 0.$$

Further, evaluating this difference:

$$(10) \quad \delta^* I = \int_{\Omega} (\delta^* \mathcal{L} + (\mathcal{L} \delta\zeta^\nu)_{,\nu}) d_4 x,$$

or explicitly:

$$(11) \quad \delta^* I = \int_{\Omega} \left(\frac{\partial \mathcal{L}}{\partial g^{\alpha\beta}} \delta^* g^{\alpha\beta} + \frac{\partial \mathcal{L}}{\partial g^{\alpha\beta}{}_{,\nu}} \delta^* g^{\alpha\beta}{}_{,\nu} + \frac{\partial \mathcal{L}}{\partial \psi_A} \delta^* \psi_A + \frac{\partial \mathcal{L}}{\partial \psi_{A,\nu}} \delta^* \psi_{A,\nu} + (\mathcal{L} \delta\zeta^\nu)_{,\nu} \right) d_4 x.$$

3. The field $\delta\zeta^\nu$ being arbitrary, let us assume that $\delta\zeta^\nu$ and its first derivatives vanish on the boundary of Ω . Thus, bearing in mind that $\delta^* \varphi_{,\nu} = (\delta^* \varphi)_{,\nu}$, and integrating by parts we obtain from (11):

$$(12) \quad \delta^* I = \int_{\Omega} \sqrt{-g} \left(\frac{1}{2} T_{\alpha\beta} \delta^* g^{\alpha\beta} + L^A \delta^* \psi_A \right) d_4 x;$$

*) More precisely, $\delta^* \psi_A$ denotes the principal part of this difference. $\delta^* \psi_A$ differs only in sign from the Lie differential treated in [4].

$\delta^* g^{\alpha\beta}$ can be easily computed:

$$(13) \quad \delta^* g^{\alpha\beta} = \delta \zeta^{\alpha;\beta} + \delta \zeta^{\beta;\alpha}.$$

Substituting (7) and (13) in (12) and integrating by parts we have:

$$\delta^* I = - \int_{\Omega} \sqrt{-g} \delta \zeta^{\alpha} \left(T_{\alpha\beta;\beta} + L^A \psi_{A;\alpha} + \frac{1}{\sqrt{-g}} (\sqrt{-g} L^A F_{A\alpha}{}^{B\beta} \psi_B)_{,\beta} \right) d_4 x \equiv 0$$

and, by virtue of (8) and the arbitrariness of Ω and $\delta \zeta^{\alpha}$:

$$(14) \quad (T^{\beta}_{\alpha} + L^A F_{A\alpha}{}^{B\beta} \psi_B)_{;\beta} + L^A \psi_{A;\alpha} \equiv 0.$$

This is a set of "strong" equations, holding independently, whether the field equations (4) are satisfied or not. When $L^A_{;\alpha} = 0$, we obtain from (14) a "weak" conservation law for the energy-momentum tensor: $T_{\alpha;\beta} = 0$.

4. Introducing the canonical (in general non-symmetric) energy-momentum density tensor $t^{\beta}_{\alpha} \stackrel{\text{def.}}{=} \delta^{\beta}_{\alpha} L + \frac{\partial L}{\partial \psi_{A;\beta}} \psi_{A;\alpha}$, we can, by a method developed by Hilbert [7] or de Wet [8], obtain a general relation between T^{β}_{α} and t^{β}_{α} . We give the formula in question for the case when L does not depend on $g_{\mu\nu,\sigma}$ (this assumption will be made throughout section 4):

$$(15) \quad T^{\beta}_{\alpha} \equiv t^{\beta}_{\alpha} - F_{A\alpha}{}^{B\beta} \left(\frac{\partial L}{\partial \psi_A} \psi_B + \frac{\partial L}{\partial \psi_{A;\nu}} \psi_{B;\nu} \right).$$

It is worth observing that in a non-Euclidean space-time, the canonical energy-momentum tensor is divergenceless even when (4) holds, viz.:

$$(16) \quad t^{\beta}_{\alpha;\beta} \equiv -L^A \psi_{A;\alpha} - \frac{\partial L}{\partial \psi_{A;\beta}} \psi_B F_{A\alpha}{}^{B\beta} R^{\mu}_{\nu\alpha\beta}.$$

When the field equations (4) are satisfied, we can write (15) in the form:

$$T^{\beta}_{\alpha} = t^{\beta}_{\alpha} + S^{\beta\nu}_{\alpha;\nu}, \text{ where}$$

$$S^{\alpha\beta\nu}_{\text{def.}} = F_{A\alpha}{}^{B\beta} \psi_B \frac{\partial L}{\partial \psi_{A;\nu}} = -S^{\nu\beta}_{\alpha}.$$

The tensor $s^{\alpha\beta\nu}_{\text{def.}} = S^{\beta\alpha\nu} - S^{\alpha\beta\nu}$ corresponds — when the V_4 is flat — to the spin angular momentum density tensor [5].

5. The problem of the motion of bodies in the general relativity theory has been investigated by many authors. Einstein and Infeld represented the bodies as singularities of the field (see for example [9]). In this paper, we assume that matter is characterised by a scalar density of mass $\varrho(x^\nu)$ and a velocity field $u^a = dx^a/ds$ (perfect fluid without pressure).

We assume that the left hand side of the gravitational equations can be obtained by varying the integral

$$J = \int_{\Omega} K \sqrt{-g} d_4 x.$$

K is an invariant built from $g_{\mu\nu}$ and their derivatives. The gravitational field equations are

$$(17) \quad K_{\alpha\beta} \stackrel{\text{def.}}{=} \frac{\delta J}{\delta g^{\alpha\beta}} = \varrho u_{\alpha} u_{\beta} + T_{\alpha\beta}.$$

For $K = zR$, we have $K_{\alpha\beta} = z(R_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} R)$, and (17) becomes Einstein's equation. In a manner similar to that used in section 3, we get $K^{\alpha\beta}_{;\beta} = 0$.

Equation (17) can be integrated only when

$$(18) \quad (\varrho u^{\alpha} u^{\beta} + T^{\alpha\beta})_{;\beta} = 0.$$

Assuming $L^A = j^A$, we obtain by virtue of (14), and $u^{\alpha}_{;\beta} u^{\beta} = Du^{\alpha}/ds$:

$$(19) \quad \varrho \frac{Du_{\alpha}}{ds} + u_{\alpha} (\varrho u^{\beta})_{;\beta} = j^A \psi_{A;\alpha} + F_{A\alpha}{}^{B\beta} (j^A \psi_B)_{;\beta}.$$

Transvecting (19) with u^{α} , we get the law of "conservation of mass":

$$(20) \quad (\varrho u^{\beta})_{;\beta} = u^{\alpha} (j^A \psi_{A;\alpha} + F_{A\alpha}{}^{B\beta} (j^A \psi_B)_{;\beta}).$$

From (19) and (20) we have the *equations of motion*:

$$(21) \quad \varrho \frac{Du^{\alpha}}{ds} = (g^{\alpha\nu} - u^{\alpha} u^{\nu}) ((j^A \psi_B)_{;\beta} F_{A\nu}{}^{B\beta} + j^A \psi_{A;\nu}).$$

The equations of motion appear as necessary *integrability conditions* for the gravitational field equations.

The author wishes to thank R. Gajewski for valuable discussions and collaboration in writing section 4.

INSTITUTE OF PHYSICS, POLISH ACADEMY OF SCIENCES

REFERENCES

- [1] L. Infeld, *Equations of motion for linear field theories*, Bull. Acad. Polon. Sci. Cl. III (1955), 213.
- [2] E. Noether, Nachr. Ges. Göttingen (1918), 235.
- [3] E. L. Hill, Rev. Mod. Phys. **23** (1951), 253.
- [4] J. A. Schouten, *Ricci Calculus*, 2nd edition, Berlin 1954.
- [5] F. J. Belinfante, Physica **7** (1940), 449.
- [6] P. G. Bergmann, Phys. Rev. **75** (1949), 680.
- [7] D. Hilbert, Math. Ann. **92** (1924), 1.
- [8] J. S. de Wet, Proc. Cambr. Phil. Soc. **43** (1947), 511.
- [9] L. Infeld, Acta Phys. Polon. **13** (1954), 187.

Killing's Equations and Conservation Theorems

by

A. TRAUTMAN

Presented by L. INFELD on August 14, 1956

1. The principle of general covariance leads, for a given field theory, to four differential identities [1]. When the field equations (I. 4)* are fulfilled, these identities yield four conservation laws

$$(1) \quad T_{\alpha;\beta}^{\beta} = 0.$$

It is known, however, that a Lorentz-covariant theory in a flat space-time possesses 10 conservation laws [2], [3]. A theory which is general-covariant is also Lorentz-covariant, but no angular momentum tensor exists in a Riemannian space-time (apart from special cases).

Thus we see that the properties of the space-time continuum are essential for deriving the conservation theorems. The purpose of this note is to reveal some of the relations between covariance, space-time structure and conservation laws. The notation of [1] will be used in this paper.

2. We shall deal with field theories, the equations of which can be obtained from (I. 1, 2), where

$$(2) \quad L = L(g_{\mu\nu}, \psi_A, \psi_{A,\nu}).$$

In general, the behaviour of the Lagrangian density L can be investigated under certain groups of transformations of the

(a) independent variables $x^\nu \rightarrow x^{\nu'}$;†

(b) metric tensor $g_{\alpha\beta} \rightarrow g'_{\alpha\beta}$;

(c) field variables $\psi_A \rightarrow \psi'_A$ ("gauge transformations").‡

We have dealt with the transformations (a) in [1]. This group will be investigated from another point of view in sections 3 and 5. The gauge transformations and related properties of the field equations are examined in [2] and [3].

*) The formulae from [1] we shall refer to as I.

Let us fix our attention on the group (b). The simplest transformations of the metric are the *conformal transformations*:

$$(3) \quad g_{\alpha\beta} \rightarrow 'g_{\alpha\beta} = \sigma g_{\alpha\beta}, \quad \text{where} \quad \sigma(x_\mu^v) \text{ is a scalar.}$$

Taking the infinitesimal conformal transformation

$$\delta g^{\alpha\beta} = 'g^{\alpha\beta} - g^{\alpha\beta} = g^{\alpha\beta} \delta\eta,$$

we obtain for the variation of $\mathfrak{L}$ (cf. (I. 5) and (2)):

$$\delta\mathfrak{L} = \frac{\partial\mathfrak{L}}{\partial g^{\alpha\beta}} g^{\alpha\beta} \delta\eta = \frac{1}{2} \sqrt{-g} T_{\alpha\beta} g^{\alpha\beta} \delta\eta = \frac{1}{2} \mathfrak{T} \delta\eta.$$

Thus, a necessary and sufficient condition for the field equations derived from a Lagrangian density (2) to be conform-invariant is

$$(4) \quad T^\alpha_\alpha = T = 0.$$

Equation (4) holds for the electromagnetic field [4] which is intimately connected with the vanishing of the rest-mass of photons.

3. Equation (1) does not allow a direct physical interpretation (in a curved space-time), because of the term $\Gamma^\nu_{\alpha\beta} \mathfrak{T}^\beta_\nu$ in $\mathfrak{T}_{\alpha;\beta}^\beta$.

For physical interpretation, conservation laws in the form of an ordinary divergence are of interest. Our purpose is to give a general method for constructing such laws.

Let us rewrite equation (I. 11) profiting from (2):

$$(5) \quad \delta^* I = \int_{\Omega} \left(\frac{\partial\mathfrak{L}}{\partial g^{\alpha\beta}} \delta^* g^{\alpha\beta} + \frac{\partial\mathfrak{L}}{\partial\psi_A} \delta^* \psi_A + \frac{\partial\mathfrak{L}}{\partial\psi_{A,\nu}} \delta^* \psi_{A,\nu} + (\mathfrak{L} \delta\zeta^\nu)_{,\nu} \right) d_4 x \equiv 0.$$

Let us now assume that in the Riemannian space-time V_4 there exists a vector field $\delta\zeta^\nu$ fulfilling the *Killing equations*:

$$(6) \quad \delta^* g^{\alpha\beta} \equiv \delta\zeta^\alpha{}_{;\beta} + \delta\zeta^\beta{}_{;\alpha} = 0.$$

Each solution of (6) generates a one-parameter group of transformations leaving the metric $g^{\alpha\beta}$ invariant in form. These transformations are called *motions* [5] and are a generalisation of Lorentz transformations in a pseudo-Euclidean space. The number of independent solutions of (6) in a V_4 is ≤ 10 . Let us by $\delta\zeta^\nu$ ($i=1,2,\dots,k \leq 10$) denote these solutions, then, writing $\delta^*_i \varphi$ for variations corresponding to $\delta\zeta^\nu_i$, we obtain from (5):

$$\delta^* I = \int_{\Omega} \left(\frac{\partial\mathfrak{L}}{\partial\psi_{A,i}} \delta^* \psi_A + \frac{\partial\mathfrak{L}}{\partial\psi_{A,\nu,i}} \delta^* \psi_{A,\nu} + (\mathfrak{L} \delta\zeta^\nu_i)_{,\nu} \right) d_4 x \equiv 0.$$

If the field equations (I. 4) are satisfied, we get

$$(7) \quad \delta^* I = \int_{\Omega} \left(\frac{\partial\mathfrak{L}}{\partial\psi_{A,\nu,i}} \delta^* \psi_A + \mathfrak{L} \delta\zeta^\nu_i \right)_{,\nu} d_4 x = 0.$$

Owing to the arbitrariness of Ω , (7) yields

$$(8) \quad \left(\frac{\partial \Omega}{\partial \psi_{A,v} i} \delta^* \psi_A + \Omega \delta \zeta^v \right)_{,v} = 0, \quad i=1, \dots, k.$$

Equation (7) when applied to a region bounded by two space-like hyper-surfaces, gives:

$$(9) \quad \int_{\infty} \left(\frac{\partial \Omega}{\partial \psi_{A,v} i} \delta^* \psi_A + \Omega \delta \zeta^v \right) dS_v = \text{const}, \quad i=1, \dots, k$$

(we have assumed the vanishing of the field in space infinity). The assumption that Ω does not depend on $g_{a\beta, v}$ played no role in this section.

4. EXAMPLES. a) When the space-time is *flat* (pseudo-Euclidean, R_4), there exist 10 independent solutions of (6). In a Galilean (pseudo-Cartesian) co-ordinate systems, we can take as solutions of (6):

$$(10) \quad \delta \zeta^v = \delta a^v, \quad \delta \zeta^v = x^\mu \delta \omega^\nu{}_\mu.$$

δa^v and $\delta \omega^{\mu\nu}$ are constant, arbitrary vector and bivector respectively. Writing (8) explicitly for $\delta \zeta^v$ given by (10), we get 4 laws of conservation of energy and momentum and 6 laws of conservation of angular momentum [3]:

$$(11) \quad t_{\mu, v}^v = 0, \quad \text{where} \quad t_\mu^v = -\delta_\mu^v L + \frac{\partial L}{\partial \psi_{A, v}} \psi_{A, \mu},$$

$$(12) \quad J_{a\beta, v}^v = 0, \quad \text{where} \quad J_{a\beta}^v = x_{[a} t_{\beta]}^v + \frac{\partial L}{\partial \psi_{A, v}} \psi_{B A} F_{[a}{}^B{}_{\beta]}.$$

b) We call a space-time *stationary*, when there exists a co-ordinate system in which $g_{a\beta, 0} = 0$. The vector field $\delta \zeta^v$ with components $\delta \zeta^0 = 1$, $\delta \zeta^k = 0$ fulfills (6), and we obtain from (8) a law of "conservation of energy":

$$\left(\frac{\partial \Omega}{\partial \psi_{A, v}} \psi_{A, 0} \right)_{, 0} = \Omega_{, 0}.$$

c) In a V_4 of *constant curvature* equations, (6) have 10 independent solutions and (8) gives 10 conservation laws.

5. Co-ordinate transformations generated by solutions of

$$(13) \quad \delta^* g^{a\beta} \equiv \delta \zeta^{\alpha; \beta} + \delta \zeta^{\beta; \alpha} = 2g^{a\beta} \delta \chi, \quad (\delta \chi = \text{scalar})$$

are called *conformal co-ordinate transformations* (not to be confused with conformal transformations of metric, section 2). They constitute a generalisation of motions ($\delta \chi = 0$). For a solution $\delta \zeta^v$ of (13), equation (5)

becomes:

$$\delta^* I = \int_{\Omega} \left(\mathfrak{T} \delta \chi + \frac{\partial \Omega}{\partial \psi_{A i}} \delta^* \psi_A + \frac{\partial \Omega}{\partial \psi_{A, v} i} \delta^* \psi_{A, v} + (\Omega \delta \zeta^v)_{, v} \right) d_4 x = 0.$$

When $L^A = 0$ we get:

$$\delta^* I = \int_{\Omega} \left(\mathfrak{I} \delta \chi + \left(\frac{\partial \mathfrak{Q}}{\partial \psi_{A,\nu i}} \delta^* \psi_A + \mathfrak{Q} \delta \zeta^{\nu} \right)_{,i} \right) d_4 x = 0$$

or:

$$(14) \quad \left(\frac{\partial \mathfrak{Q}}{\partial \psi_{A,\nu i}} \delta^* \psi_A + \mathfrak{Q} \delta \zeta^{\nu} \right)_{,i} = - \mathfrak{I} \delta \chi.$$

For $T = 0$ (conform-invariant field theory), (14) becomes identical in form with (8).

As an example, we can take a flat space-time and the Maxwell field [4]. Equations (13) in R_4 have 15 independent solutions: 10 motions (10) and 5 infinitesimal conformal transformations which are not motions, e. g.:

$$\delta \zeta^{\nu} = x^{\nu} \delta a, \quad \delta \zeta^{\nu} = 2x^{\nu} x^{\mu} \delta a_{\mu} - x_{\mu} x^{\mu} \delta a^{\nu} \quad (\delta a, \delta a^{\nu} = \text{const}).$$

There is no difficulty in writing corresponding conservation laws.

The author would like to express his gratitude to Professors L. Infeld and J. Plebański for their kind interest in this work.

INSTITUTE OF PHYSICS, POLISH ACADEMY OF SCIENCES

REFERENCES

- [1] A. Trautman, *On the conservation theorems and equations of motion in co-variant field theories*, in this issue p. 677.
- [2] W. Pauli, *Rev. Mod. Phys.* **13** (1941), 203.
- [3] E. M. Corson, *Introduction to tensors, spinors and relativistic wave-equations*, London, 1954.
- [4] E. Bessel-Hagen, *Math. Ann.* **84** (1921), 258.
- [5] L. P. Eisenhart, *Riemannian geometry*, Princeton 1949.

Über die Bewegungsgleichungen des Pol-Dipol-Teilchens

VON

K. NAGY

Vorgelegt von L. INFELD am 8. September 1956

Mit den Bewegungsgleichungen des Pol-Dipol-Teilchens haben sich mehrere Forscher von verschiedenen Gesichtspunkten aus beschäftigt [1]—[7]. Unser Ziel ist die einfache Aufstellung der Bewegungsgleichungen des Pol-Dipol-Teilchens auf Grund der Infeldschen Methode, die für Pol-Teilchen ausgearbeitet wurde [8].

Wir betrachten das Teilchen als ein Probeteilchen. Den Spin des Teilchens beschreiben wir durch einen antisymmetrischen Tensor $S_{\alpha\beta}$.

$$(1) \quad S_{23} = \sigma_x; \quad S_{31} = \sigma_y; \quad S_{12} = \sigma_z.$$

$\sigma(\sigma_x, \sigma_y, \sigma_z)$ bedeuten hier den Spinvektor. Wir nehmen an, dass die $S_{\alpha\beta}$ in dem Koordinatensystem in dem das Teilchen ruht gleich Null sind. Diese Bedingung kann man in der folgenden kovarianten Form anschreiben:

$$(2) \quad S_{\alpha\beta} u_\beta = 0,$$

wo u_β die Vierergeschwindigkeit des Teilchens bedeutet.

Die Bewegungsgleichungen kann man aus den Feldgleichungen nur in nichtlinearen Feldtheorien herleiten. Nach dem Gedankengang von L. Infeld nehmen wir zu den Gleichungen des linearen Feldes (z. B. des elektro-magnetischen-, oder Mesonfeldes) noch die Gleichungen des Gravitationsfeldes hinzu, und bekommen so ein nichtlineares Gleichungssystem, aus dem man die Bewegungsgleichungen herleiten kann. Gehen wir endlich zu der speziellen Relativitätstheorie durch den Grenzübergang $g_{\alpha\beta} \rightarrow \sigma_{\alpha\beta}$, $x \rightarrow 0$ (x bedeutet hier die Gravitationskonstante) über, so bekommen wir die Bewegungsgleichungen in den linearen Feldtheorien:

$$(3) \quad \frac{dt}{ds} \int T_{\alpha\beta,\beta} dv = 0.$$

Der Energie-Impuls-Tensor $T_{\alpha\beta}$ ist

$$(4) \quad T_{\alpha\beta} = t_{\alpha\beta} + E_{\alpha\beta},$$

wo $t_{\alpha\beta}$ den kinematischen Energie-Impuls-Tensor des Pol-Dipol-Teilchens bedeutet und $E_{\alpha\beta}$ der Energie-Impuls-Tensor des linearen Feldes ist. (Der Tensor $E_{\alpha\beta}$ enthält auch den aus der Wechselwirkung zwischen Feld und Teilchen stammenden Energie-Impuls-Tensor).

Betrachten wir hier nur den Fall der kräftefreien Bewegung: $E_{\alpha\beta} = 0$. Nehmen wir für den Tensor $t_{\alpha\beta}$ den folgenden Ausdruck an [9]:

$$(5) \quad t_{\alpha\beta} = \int (m_{\alpha\beta} + m_{\alpha\beta\gamma\zeta\gamma}) \delta_{(4)} ds,$$

wo

$$(6) \quad m_{\alpha\beta} = m_0 u_\alpha u_\beta - \frac{1}{2c^2} (\dot{S}_{\alpha\nu} u_\nu u_\beta + \dot{S}_{\beta\nu} u_\nu u_\alpha),$$

$$(7) \quad m_{\alpha\beta\gamma\zeta\gamma} = -\frac{1}{2} (\dot{S}_{\gamma\alpha} u_\beta + \dot{S}_{\gamma\beta} u_\alpha)$$

ist. m_0 ist hier die Ruhmasse des Teilchens, der Punkt bedeutet die Ableitung nach der Eigenzeit s . $\delta_{(4)}$ ist die vierdimensionale Diracsche δ -Funktion.

Die Bewegungsgleichungen sind auf Grund der Beziehungen (3), (4), (5), (6) und (7):

$$(8) \quad \frac{dP_\alpha}{ds} = 0,$$

wo

$$(9) \quad P_\alpha = m_0 u_\alpha - \frac{1}{c^2} \dot{S}_{\alpha\nu} u_\nu + \frac{1}{2c^2} \left(\dot{S}_{\alpha\nu} u_\nu - \frac{\dot{S}_{4\nu}}{u_4} u_\nu u_\alpha + c^2 \frac{\dot{S}_{\alpha\gamma}}{u_\gamma} \right)$$

ist. (Später werden wir zeigen, dass der in den Klammern stehende Ausdruck in Gl. (9) verschwindet).

Betrachten wir jetzt den Drehimpuls des Teilchens:

$$(10) \quad N_{\alpha\beta} = \frac{1}{ic} \int (x_\alpha t_{\beta 4} - x_\beta t_{\alpha 4}) dv.$$

Bei Berücksichtigung der Beziehungen (5), (6), und (7) kann man schreiben

$$(11) \quad N_{\alpha\beta} = \xi_\alpha P_\beta - \xi_\beta P_\alpha + S_{\alpha\beta}.$$

($\xi_\alpha(s)$ sind die Koordinaten des Teilchens). Der gesamte Drehimpuls ist gleich: Bahndrehimpuls $\xi_\alpha P_\beta - \xi_\beta P_\alpha$ plus innerer Drehimpuls (Spin) $S_{\alpha\beta}$. Aus der Gleichung $t_{\alpha\beta} = 0$ folgt, dass $N_{\alpha\beta}$ eine Konstante ist. Dementsprechend ist also

$$(12) \quad S_{\alpha\beta} + u_\alpha P_\beta - u_\beta P_\alpha = 0.$$

Durch Überschieben von u_β erhält man wegen $u_\nu u_\nu = c^2$, dass die Klammer in Gl. (9) gleich Null ist.

Die Bewegungsgleichungen des Pol-Dipol-Teilchens sind endlich:

$$(8') \quad \frac{d}{ds} \left(m_0 u_\alpha - \frac{1}{c^2} \dot{S}_{\alpha\nu} u_\nu \right) = 0.$$

Für den Spintensor $S_{\alpha\beta}$ bekommt man die folgenden Bewegungsgleichungen:

$$(12') \quad \dot{S}_{\alpha\beta} + \frac{1}{c^2} (\dot{S}_{\alpha\nu} u_\nu u_\beta - \dot{S}_{\beta\nu} u_\nu u_\alpha) = 0.$$

Aus der Gl. (8) bekommen wir durch Überschieben von u_α , dass die Ruhmasse des Teilchens im kräftefreien Falle eine Konstante ist.

Eine ausführliche Mitteilung wird in den *Acta Physica Hungarica* erscheinen.

INSTITUT FÜR THEORETISCHE PHYSIK DER ROLAND EÖTVÖS UNIVERSITÄT IN
BUDAPEST

SCHRIFTTUM

- [1] J. Frenkel, Z. Phys. **37**, (1926), 572.
- [2] M. Mathisson, Acta Phys. Polon. **6** (1937), 167.
- [3] J. Lubański, Acta Phys. Polon. **6**, (1937) 356.
- [4] H. Hönl u. A. Papapetrou, Z. Phys. **112** (1939), 512.
- [5] — Z. Phys. **114** (1939), 478.
- [6] J. Weyssenhoff u. A. Raabe, Acta Phys. Polon. **9** (1946).
- [7] A. Papapetrou u. W. Urich, Z. Naturforsch. **10a**, (1955), 109.
- [8] L. Infeld, Bull. Acad. Polon. Cl. III **3** (1955), 213.
- [9] H. Hönl, Ergebn. exact. Naturw. **26** (1952), 291.

On Modified Dirac δ -Functions

by

L. INFELD and J. PLEBAŃSKI

Presented by L. INFELD on August 25, 1956

1. Introduction. Dirac's δ -functions play an important role in theoretical physics, especially in classical and quantum field theories. Mathematicians find the use of these δ -functions objectionable, while physicists find the use of " δ distributions" much too abstract.

The present note is written by physicists and we hope it will be of use to physicists. Its aim is to show how, by narrowing the concept of Dirac's δ -functions, we can enrich them with properties which are useful for applications, especially when dealing with some infinities. Indeed, this paper arose from the necessity of applying such δ -functions to the problem of motion in the theory of gravitation.

2. On the Introduction of Dirac's δ -functions. One usually introduces the function $\delta(\vec{x})$ *) in one of three possible ways: the axiomatic method (A); the method of Fourier-transforms (F); the realistic method (R). Let us briefly recapitulate these methods.

The (A) method defines $\delta(\vec{x})$ in the following way:

(A₁) — $\delta(\vec{x})$ can be formally differentiated — that is, there exist deriva-

tives $\frac{\partial^{a+b+c}}{\partial x_1^a \partial x_2^b \partial x_3^c} \delta(\vec{x})$.

(A₂) — $\delta(\vec{x}) = 0$ if $\vec{x} \neq 0$; $\delta(0) = +\infty$.

(A₃) — For an arbitrary region $\Omega(\vec{x}_0)$ containing $\vec{x}_0$ as an internal point and for an arbitrary continuous function $f(\vec{x})$ we have

$$\int_{\Omega(\vec{x}_0)} d_{(3)}x \delta(\vec{x} - \vec{x}_0) f(\vec{x}) = f(\vec{x}_0).$$

*) We shall deal here only with three-dimensional δ 's. The theory can easily be translated to one or more dimensional δ 's.

The (F) method defines $\delta(\vec{x})$ by the equation

$$(F) \quad \delta(\vec{x}' - \vec{x}'') = \sum_{\text{over spectrum } \lambda} \Psi_{\lambda}^*(\vec{x}') \Psi_{\lambda}(\vec{x}''),$$

where $\Psi_{\lambda}(\vec{x})$ are the eigenfunctions satisfying the equation $\Lambda \Psi_{\lambda}(\vec{x}) = \lambda \Psi_{\lambda}(\vec{x})$ with corresponding boundary conditions, where Λ is a complete set of real operators. When, for example, $\Psi_{\lambda}(\vec{x}) = \frac{1}{(\sqrt{2\pi})^3} e^{i\vec{k}\vec{x}}$, being the eigenfunctions of the operators $-i \frac{\partial}{\partial x_a}$, we have

$$\delta(\vec{x}) = \frac{1}{(2\pi)^3} \int d_3 k e^{-i\vec{k}\vec{x}}.$$

For $\delta(\vec{x})$ defined by means of (F), it is assumed that differentiation (and integration) can replace the summation over the λ spectrum. Then we "deduce" (A₁₋₃) from the properties of the complete set of eigenfunctions.

In the (R) method — in which we shall be most interested here — $\delta(\vec{x})$ is defined as the limit of a sequence of functions satisfying certain conditions. Such a procedure can be connected with a certain concrete picture: $\delta(\vec{x})$ is the limit of a density distribution of a unit mass if the mass is concentrated at one point.

We shall now try to connect a realistic picture with the definition of $\delta_n(\vec{x})$, leading to $\delta(\vec{x})$ in the following way:

We shall say that a sequence $\{\delta_n(\vec{x})\} \in D$, by which we mean that $\delta_n(\vec{x})$ belongs to the class of sequences of models of δ -functions if:

- (R₁) The $\delta_n(\vec{x})$ are functions determined for every $\vec{x}$ and of class C_{∞} (meaning they all have derivatives);
- (R₂) for every $\vec{x} \neq 0$ there exists $\lim_{n \rightarrow \infty} \delta_n(\vec{x}) = 0$;
- (R₃) for every region $\Omega(\vec{x}_0)$ containing $\vec{x}_0$ as an internal point and for every continuous function $f(\vec{x})$ in $\Omega(\vec{x}_0)$, there exists a limit:

$$\lim_{n \rightarrow \infty} \int_{\Omega(\vec{x}_0)} d_3 x \delta_n(\vec{x} - \vec{x}_0) f(\vec{x}) = f(\vec{x}_0),$$

where the integral is an ordinary Riemann integral.

Indeed, this realistic method leads to the axiomatic method, especially to (A₃), if — without mathematical justification — we replace the limiting process $n \rightarrow \infty$ with the integral and denote $\lim_{n \rightarrow \infty} \delta_n(\vec{x}) = \delta(\vec{x})$.

Of course, we can give an infinite number of examples of sequences

$$\{\delta_n(\vec{x})\} \in D \quad (\text{e. g. } \delta_n(\vec{x}) = \frac{1}{(\sqrt{2\pi})^3} \frac{1}{\varepsilon_n^3} \exp \left[-\frac{1}{2} \frac{|\vec{x}|^2}{\varepsilon_n^2} \right]),$$

where ε_n is a sequence of numbers monotonically converging to zero). This realistic method is a *justification* of the axiomatic method. Indeed we could treat all calculations with δ -functions as calculations with $\delta_n(\vec{x})$ and go to the limit $n \rightarrow \infty$ at the end of these calculations. When the end result does not depend on the model used, we are justified in formally using the δ 's according to their axiomatic definition.

3. On the introduction of modified δ -functions. We shall try to narrow down the definition of Dirac's δ -function, denoting the new functions by $\hat{\delta}$. We shall start by an axiomatic definition ($\hat{A}$) as we did with the ordinary δ -functions and show later that a realistic definition ($\hat{R}$) is also possible. Thus the realistic definition will be, as before, the justification for an axiomatic definition. We are not concerned whether the number of axioms or realistic properties given here is necessary. We are only concerned to know whether they are sufficient.

Axiomatically we define $\hat{\delta}(x)$ as follows:

($\hat{A}_1$) $\hat{\delta}(x)$ can be formally differentiated up to any order.

($\hat{A}_2$) $\hat{\delta}(\vec{x}) = 0$ if $\vec{x} \neq 0$: $\hat{\delta}(\vec{x})$ can be treated as a spherically symmetric function.

($\hat{A}_3$) For an arbitrary region $\Omega(\vec{x}_0)$ containing x_0 as an internal point and for an arbitrary continuous function $f(\vec{x})$ of $\vec{x}$ in $\Omega(\vec{x}_0)$, we have

$$\int_{\Omega(\vec{x}_0)} d_3x \hat{\delta}(\vec{x} - \vec{x}_0) f(\vec{x}) = f(\vec{x}_0).$$

($\hat{A}_4$) The integrals

$$\int_{\Omega(0)} d_3x x \frac{\hat{\delta}(\vec{x})}{|\vec{x}|^p}, \quad p = 1, 2, \dots, k,$$

have a definite meaning and they fulfill the condition (for every region of integration $\Omega(0)$)

$$\int_{\Omega(0)} d_3x \hat{\delta}(\vec{x}) \frac{1}{|\vec{x}|^p} = 0, \quad p = 1, 2, \dots, k.$$

And now we shall justify ($\hat{A}_4$). Indeed it seems almost shocking to find ($\hat{A}_4$) among the axioms. Yet it can be shown that a class of models exists which satisfies all these conditions. This means that we can choose a $\{\hat{\delta}_n(\vec{x})\} \in \hat{D}_k$ such that:

($\hat{R}_1$) $\{\hat{\delta}_n(\vec{x})\} \in D$.

($\hat{R}_2$) For every n $\hat{\delta}_n(\vec{x}) = 0$ ($|\vec{x}|^{k-2}$), for $|\vec{x}| \ll 1$, where k is an integer characterising the class $\hat{D}_k$. This means

$$\lim_{|\vec{x}| \rightarrow 0} |\vec{x}|^{-k+2+\varepsilon} \hat{\delta}_n(\vec{x}) = 0, \quad \text{for } \varepsilon > 0.$$

($\hat{R}_3$) For an arbitrary $\Omega(0)$ surrounding $\vec{x}=0$, the limits

$\lim_{n \rightarrow \infty} \int_{\Omega(0)} d_3 x \frac{\hat{\delta}_n(\vec{x})}{|\vec{x}|^p} x=0, p=1, 2, \dots, k$, exist, where the integrals have to be understood in the Riemannian sense. Their existence is implied by ($\hat{R}_2$).

To construct $\hat{\delta}_n(\vec{x})$, we proceed as follows:

(C_1) We choose $\Delta(z) \in C_\infty$ for $z \geq 0$, which approaches zero, at least exponentially if $z \rightarrow \infty$. For $z \rightarrow 0$, $\Delta(z) = 0(z^{k-2})$. Furthermore:

$$4\pi \int_0^\infty dz z^2 \Delta(z) = 1 \quad \left(\text{e. g. } \Delta(z) = \frac{1}{2\pi} \frac{Z^{k-2}}{2^{\frac{k+1}{2}} \Gamma\left(\frac{k+1}{2}\right)} e^{-\frac{1}{2}z^2} \right).$$

(C_2) For Dirac's δ -function we choose $\delta_\varepsilon(\vec{x}) = \frac{1}{\varepsilon^3} \Delta\left(\frac{|\vec{x}|}{\varepsilon}\right)$. (It is easy to see that this is one of the realistic models of Dirac's δ -functions).

(C_3) We define: $\hat{\delta}_n(\vec{x}) = \int_0^1 du I_k(u) \delta_{\varepsilon_n u}(\vec{x})$; ($\varepsilon_n \rightarrow 0$), where $I_k(u) =$

$(-)^k \frac{(2k+1)!}{(k!)^3} u^k \left(\frac{d}{du}\right)^k u^k (1-u)^k$. The functions $\left(\frac{d}{du}\right)^k u^k (1-u)^k$ are Jacobian polynomials of order k with the two parameters equal to one. Indeed, it is easy to show that such functions are a narrow class of Dirac's δ_n functions since they satisfy R . This is obvious for R_1 and R_2 . As for R_3 , we see, because of the normalization of I_k — that is. because

$$\int_0^1 du I_k(u) = 1,$$

that

$$\lim_{n \rightarrow \infty} \int_{\Omega(0)} d_3 x \hat{\delta}_n(\vec{x}) = 1.$$

Now, as to the desired property ($\hat{A}_4$):

$$\begin{aligned} \lim_{n \rightarrow \infty} \int_{\Omega(0)} d_3 x \frac{\hat{\delta}_n(\vec{x})}{|\vec{x}|^p} &= \lim_{n \rightarrow \infty} \int_0^\infty d_3 x \frac{\hat{\delta}_n(\vec{x})}{|\vec{x}|^p} = \lim_{n \rightarrow \infty} \int_0^\infty d_3 x \frac{1}{|\vec{x}|^p} \int_0^1 du I_k(u) \delta_{\varepsilon_n u}(\vec{x}) = \\ &= \lim_{n \rightarrow \infty} \int_0^1 du I_k(u) \int_0^\infty d_3 x \frac{1}{|\vec{x}|^p} \frac{1}{(\varepsilon_n u)^3} \Delta\left(\frac{|\vec{x}|}{\varepsilon_n u}\right) = \lim_{n \rightarrow \infty} \varepsilon_n^{-p} \int_0^1 du u^{-p} I_k(u) \int_0^\infty d_3 z \frac{\Delta\left(\frac{|\vec{z}|}{u}\right)}{|\vec{z}|^p} = 0. \end{aligned}$$

The last equation is true because

$$\int_0^1 du u^{k-p} \left(\frac{d}{du}\right)^k u^k (1-u)^k = 0, \quad p=1, 2, \dots, k.$$

Now — a short comment on this proof: We may change the sequence of integrals with respect to $d_3 \hat{x}$ and $d\hat{u}$ because the integral $\int_{\infty} d_3 z \frac{1}{|\hat{z}|^p} A(|\hat{z}|)$ exists, that is because $A(z) = O(z^{k-2})$, $z \ll 1$. In the last step of our proof we changed the variables $\hat{x}/\varepsilon_n u$ to $\hat{z}$.

One more comment: The condition $A(z) = O(z^{k-2})$ is not necessary. But the proofs become much simpler if this condition is assumed.

INSTITUTE OF PHYSICS, POLISH ACADEMY OF SCIENCES
INSTITUTE OF THEORETICAL PHYSICS, WARSAW UNIVERSITY

Azeotropic Depression as a Function of the Composition of Series of Ternary Positive Azeotropes

by

W. ŚWIĘTOSŁAWSKI and W. MALESIŃSKI

Presented by W. ŚWIĘTOSŁAWSKI on August 18, 1956

1. Introductory remarks. The relation between the composition and azeotropic depression in a series of binary positive homoazeotropes was given in two recent papers [1], [2]. It was shown on the basis of formulae determining the activity coefficients, valid in the case of regular solutions, that the azeotropic points in a series of binary azeotropes (A, H_i) is related as follows:

$$(1) \quad T_A - T_{Az} = Z_d \phi_{H_i}^2,$$

T_A being the boiling temperature of the azeotropic agent A , T_{Az} — that of the azeotrope, x_{H_i} — the mole fraction of homologue H_i in the azeotropic mixture, and Z_d — the lower portion of the azeotropic range of the agent in relation to the homologous series.

If reduced values of the azeotropic depressions are introduced [2], then the azeotropic points for different agents and homologous series lie on a common parabolic curve, which, together with equation (1), agrees well with the available experimental data. Greater deviations were observed in the case of highly polar and associated azeotropic agents.

Now is to be derived the equation relating the azeotropic depression with the composition of individual ternary azeotropes. Replacing successively one of the components by its homologues, we obtain a series of ternary azeotropes and, it will be seen, the azeotropic parameters of such series are related quite analogously to equation (1).

An attempt in this direction was made by Orszagh [3], who extended Skolnik's empirical equation [4] to series of ternary systems.

We base our considerations on the relations between the boiling temperature and composition of individual ternary positive azeotropes [5], [6]. In addition, there is below a discussion of the change in the azeotropic composition in a series of azeotropes.

2. General equations. Our considerations are based on the previously derived equation [5] relating the azeotropic depression $T_2 - T_{Az}^{(1,2,3)}$ to the composition of ternary azeotropic mixture fixed by x_1 and x_3 :

$$(2) \quad T_2 - T_{Az}^{(1,2,3)} = \frac{A_{12}}{\Delta S_2} x_1^2 + \frac{A_{23}}{\Delta S_2} x_3^2,$$

in which T_2 is the boiling temperature of the main component, $T_{Az}^{(1,2,3)}$ — that of the azeotrope (1,2,3), ΔS_2 is the vaporisation entropy of component 2, and A_{12} and A_{23} are the regular solution constants of binary systems (1,2) and (2,3). On deriving this equation, it was assumed that the constants are related as follows:

$$(3) \quad A_{12} + A_{23} = A_{13}.$$

This empirical relation was shown to be valid for systems in which there exist no specific interactions between unlike molecules [7].

It was shown [6] that the term $\frac{A_{23}}{\Delta S_2} x_3^2$ in (2) is equal to:

$$(4) \quad (T_2 - T_{Az}^{(2,3)}) + \sqrt{\frac{T_2 - T_{Az}^{(2,3)}}{A_{23}}} \Delta S_2 (T_{Az}^{(2,3)} - T_{Az}^{(1,2,3)}) \left(1 - \frac{\Delta S_3}{\Delta S_2}\right).$$

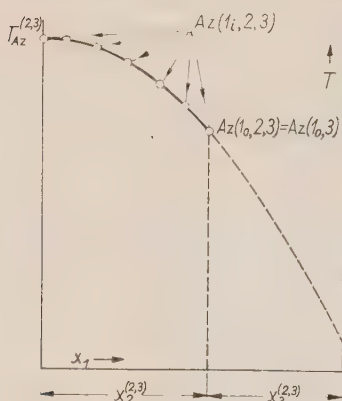
Substituting this expression into (2) we obtain:

$$(5) \quad (T_{Az}^{(2,3)} - T_{Az}^{(1,2,3)}) \left[1 - \sqrt{\frac{T_2 - T_{Az}^{(2,3)}}{A_{23}}} \Delta S_2 \left(1 - \frac{\Delta S_3}{\Delta S_2}\right)\right] = \frac{A_{12}}{\Delta S_2} x_1^2.$$

The last equation establishes the relation between the boiling temperature differences of the azeotropes (2,3) and (1,2,3) and the concentration of component 1. An analogous equation may be obtained relating the ternary azeotropic depression to the azeotrope (1,2) and x_3 .

Replacing component 1 by its homologues, i. e., taking into account a series of ternary azeotropes $(1_i, 2, 3)$, and assuming that the regular solution constants remain unchanged for all the systems with successive homologues, we may express the boiling temperature of ternary azeotropes in these series as a function of x_{1_i} only. Therefore, equation (5) may be rewritten in the form:

Fig. 1. Relation between $(T_{Az}^{(2,3)} - T_{Az}^{(1,2,3)})$ -values and x_{1_i} .



$$(6) \quad T_{Az}^{(2,3)} - T_{Az}^{(1,2,3)} = a Z_{d_{21}} x_{1_i}^2,$$

where the parameter a is the reciprocal expression in square brackets in (5) — being the consequence of the inequality of the vaporisation entropies of the components — and $Z_{d_{21}}$

is the lower portion of the azeotropic range of agent 2 in relation to series (1_i) as previously defined [1].

If $\Delta S_2 = \Delta S_3$, then $a = 1$. Even if the vaporisation entropies are different, a differs insignificantly from unity. Therefore, we may write:

$$(7) \quad T_{Az}^{(2,3)} - T_{Az}^{(1,2,3)} = Z_{d_{21}} x_{1_i}^2.$$

A graphic representation of this relation is shown in Fig. 1.

This is quite analogous to relation (1) for series of binary systems. The essential difference is that in series of ternary azeotropes, x_1 cannot undergo changes in the entire range from 0 to 1. In limit, the value of x_1 may be equal to $x_2^{(23)}$, the mole fraction of component 2 in the azeotrope (2,3) (Fig. 1).

3. Composition changes in the series of ternary positive azeotropes.

The composition changes in ternary positive azeotropes was the subject of another paper [8]. The present, more detailed discussion is based on

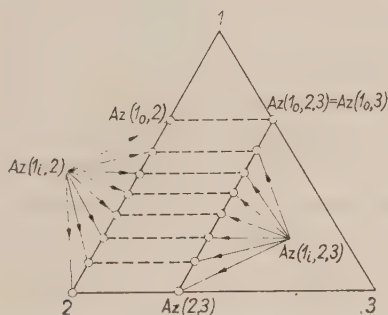


Fig. 2. Changes in azeotropic composition in the series of ternary positive azeotropes $(1_i, 2, 3)$. Case of equal vaporisation entropies of the components

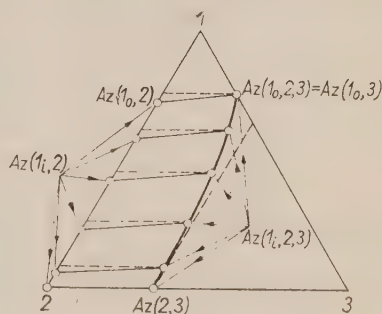


Fig. 3. Changes in azeotropic composition in the series of ternary positive azeotropes $(1_i, 2, 3)$. Case of unequal vaporisation entropies of the components

the relation referred to in our introductory remarks. The composition of an individual ternary azeotrope is as follows [5], [6]:

$$(8) \quad x_1 = x_1^{(1,2)} + \Delta x_1 \quad x_3 = x_3^{(2,3)} + \Delta x_3,$$

where $x_1^{(1,2)}$ and $x_3^{(2,3)}$ are the mole fractions of the respective components in azeotropes (1,2) and (2,3). The correction values Δx_1 and Δx_3 are due to differences in vaporisation entropies of the components. They are of the order of several mole percent [5]. Disregarding them, we have:

$$(9) \quad x_1 = x_1^{(1,2)}, \quad x_3 = x_3^{(2,3)}.$$

Assuming that in a series of ternary azeotropes, only component 1 is replaced by its homologues, we find that, in accordance with (9), the

concentration of component 3 is constant in all azeotropes of this series, being equal to that in the azeotrope (2,3). This result is, in the Gibbs triangle, represented graphically by a straight line, drawn through the azeotropic point (2,3) parallel to the 1,2-axis (Fig. 2). Successive binary

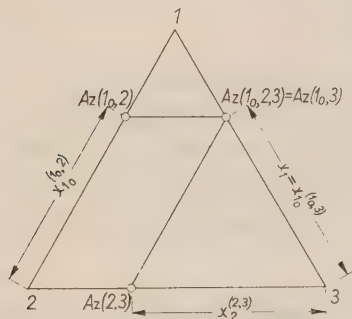


Fig. 4. Diagram of the lowest boiling tangent azeotrope ($1_0, 2, 3$) in a series of ternary azeotropes ($1_i, 2, 3$). The composition of this azeotropes is determined by

$$x_1 = x_1^{(1_0,2)} = x_1^{(1_0,3)} = x_2^{(2,3)}$$

and

$$x_3 = x_3^{(1_0,3)} = x_3^{(2,3)}$$

ternary azeotrope the following relations are valid:

$$x_1 = x_1^{(1_0,2)} = x_1^{(1_0,3)} = x_2^{(2,3)} \quad \text{and} \quad x_3 = x_3^{(1_0,3)} = x_3^{(2,3)}. \quad (\text{Fig. 4}).$$

The tangent azeotrope merging with the binary (2,3) is the highest boiling ternary azeotrope of the given series. In this case, the binary azeotrope (1,2) is a tangent azeotrope. This is clearly shown in Fig. 2. Taking into account these limiting conditions for tangent ternary azeotropes, we find that the mole fraction of the homologue (1_i) changes in the series of ternary azeotropes from 0 to $x_2^{(2,3)}$ ($0 > x_1 < x_2^{(2,3)}$).

An analogous picture of the formation of ternary azeotropes may be obtained, substituting for component 1, component 3 by its homologues, since in our considerations components 1 and 3 are equivalent.

4. The three-dimensional diagram for series of positive ternary azeotropes.

Equation (7) relating the azeotropic depression to the composition of the azeotrope, and also the relation determining the composition of ternary azeotropes, may be represented in a common diagram. Such a diagram, for the case of equal vaporisation entropies of the components, is shown in Fig. 5. The section of a parabolic curve expressed by (7) was drawn on the plane $x_3 = x_3^{(2,3)} = \text{const.}$ The section is limited by points Az (2,3) and Az ($1_0,3$), representing the highest and lowest

boiling tangent azeotropes. The projection on the triangle plane represents a diagram analogous to that shown in Fig. 2. On the three-dimensional diagram are marked the boiling temperatures of homologues (1_i) forming the lower and the upper tangent isobars with components 2 and 3: $T_{1_k}^{(2)}$, $T_{1_e}^{(2)}$ and $T_{1_k}^{(3)}$, $T_{1_e}^{(3)}$, together with the boiling temperature T_{1_0} and the isobars of 1_0 with 2 and 3. Component 1_0 is the lowest boiling component of series (1_i) forming the lowest tangent ternary azeotrope ($1_0, 2, 3$).

5. Experimental verification. No experimental data for series of ternary azeotropes are available, except those recorded by Orszagh [9] in the course of fractional distillation of a polyazeotropic mixture composed of a series of hydrocarbons from a gasoline fraction (1_i), benzene (2) and three alcohols-methanol ($3'$), isopropanol ($3''$) and isobutanol ($3'''$). Because in that work was given a characteristic only of the collected fractions, without composition and boiling temperatures of the individual azeotropes, exact verification is impossible.

Table I contains the experimental and calculated results. In successive columns are given: the boiling temperature of the fractions of ternary azeotropes $t_{Az}^{(1,2,3)}$, the depression as regards the benzene-alcohol azeotrope $t_{Az}^{(2,3)} - t_{Az}^{(1,2,3)}$, the mole fractions x_{1_i} of the hydrocarbons and x_3 — of alcohol. In the final column is shown the value of the lower azeotropic range of benzene in relation to aliphatic hydrocarbons $Z_{d_{21}}$, calculated according to (7).

The concentration x_3 of all the alcohols remains approximately constant in accordance with (9). The sign of the slight deviations from (9), i. e., the increase of x_3 with increasing boiling temperature of the ternary azeotropes, agrees with our previous considerations [5], in which the deviations Δx_1 and Δx_3 were calculated as a function of the azeotropic depression or concentration in ternary systems.

The calculated values $Z_{d_{21}}$ vary within wide limits and deviate from the known value — binary systems equal to about 18°C . of the lower portion of the azeotropic range.

Fig. 6 shows the experimental points and the theoretical curve calculated from (7) with $Z_{d_{21}} = 18^\circ\text{C}$. The limiting value of x_{1_0} is marked,

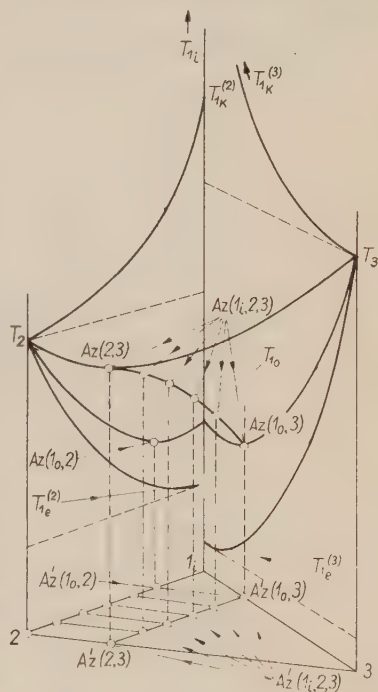


Fig. 5. Three-dimensional diagram of a series of azeotropes ($1_i, 2, 3$)

TABLE I

No.	$t_{Az}^{1,2,3}$ °C.	$t_{Az}^{2,3} - t_{Az}^{1,2,3}$ °C.	x_{1_i} %	x_3 %	$Z_{d_{21}}$
system: benzene (2) — methanol (3') — series (1 _i) $t_{Az}^{2,3} = 57.5$					
1.	53.99	3.51	38.0	57.9	27.1
2.	54.64	2.86	31.7	60.4	28.6
3.	55.34	2.16	29.4	60.8	25.0
4.	55.76	1.74	25.6	61.8	26.5
5.	56.14	1.36	21.7	61.9	28.9
6.	56.50	1.00	13.6	65.2	54.0
7.	56.89	0.61	13.6	64.5	33.0
system: benzene (2) — isopropanol (3'') — series (1 _i) $t_{Az}^{2,3} = 71.92$					
1.	60.21	11.71	66.2	30.0	26.7
2.	63.72	8.20	60.4	32.1	22.5
3.	65.66	6.26	51.4	34.7	23.7
4.	67.03	4.89	47.6	34.9	20.1
5.	68.17	3.75	42.0	36.0	21.3
6.	69.20	2.72	36.3	36.3	20.6
7.	69.95	1.97	29.9	37.2	22.0
8.	70.27	1.65	27.1	37.7	22.5
9.	70.56	1.36	22.5	39.0	26.9
10.	70.90	1.02	19.5	39.6	26.8
11.	71.12	0.80	15.1	38.9	35.0
12.	71.31	0.61	14.5	39.3	29.0
13.	71.48	0.44	9.4	40.0	49.8
14.	71.64	0.28	9.4	40.3	31.7
15.	71.76	0.16	6.9	40.6	33.6
16.	71.80	0.12	7.0	40.5	25.2
17.	71.81	0.11	5.1	41.8	42.3
system: benzene (2) — isobutanol (3''') — series (1 _i) $t_{Az}^{2,3} = 79.84$					
1.	60.43	19.41	84.4	5.7	27.3
2.	68.08	11.76	80.3	5.6	18.2
3.	70.75	9.09	75.0	6.2	16.1
4.	72.66	7.18	64.9	6.9	17.1
5.	74.12	5.72	62.9	5.7	14.4
6.	75.52	4.32	49.6	6.5	17.6
7.	76.33	3.51	44.1	6.8	18.0
8.	76.85	2.99	39.5	7.1	19.2
9.	77.22	2.62	35.4	6.6	21.0
10.	77.61	2.23	35.2	6.4	18.0
11.	77.95	1.89	28.0	7.5	24.1
12.	78.26	1.58	21.2	8.2	35.2
13.	78.63	1.21	20.1	9.0	30.0
14.	78.93	0.91	13.3	9.5	51.4
15.	79.14	0.70	14.5	9.2	33.3
16.	79.33	0.51	14.0	9.3	26.2
17.	79.54	0.30	6.8	10.1	65.0

together with the lowest boiling temperature of the lower tangent ternary azeotropes for the three series considered.

In spite of the discrepancies, due partly to deviations of alcoholic solutions from the properties of regular solutions, and partly to experimental errors, we believe that the data essentially confirm the relations derived.

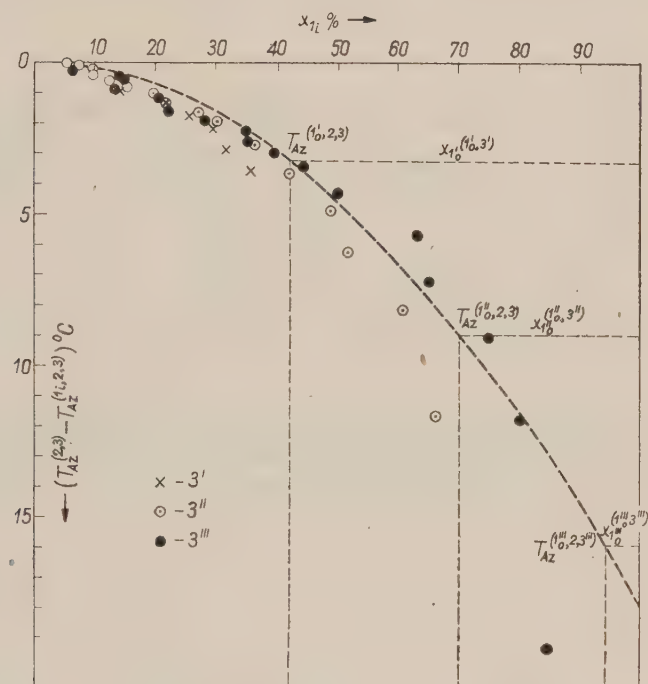


Fig. 6. Ternary azeotropic depression as regards the benzene-alcohol azeotrope as a function of the mole fraction of the hydrocarbons in the series of ternary azeotropes composed of benzene (2), methanol (3') (or isopropanol (3'') or isobutanol (3''')), and the series of aliphatic hydrocarbons (1_i)

Summary

1. Assuming that the ternary solutions considered are regular, the following equation was derived relating the boiling temperatures of a series of ternary azeotropes to their compositions:

$$(7) \quad T_{Az}^{(2,3)} - T_{Az}^{(1_i,2,3)} = Z_{d_{21}} x_{1_i}^2,$$

in which $T_{Az}^{(2,3)}$ and $T_{Az}^{(1_i,2,3)}$ denote the boiling points of the binary (2, 3) and of the successive ternary azeotropes (1_i, 2, 3), formed with the representatives (1_i) of a homologous series. $Z_{d_{21}}$ is the lower portion of the azeotropic range of the main component 2 in relation to series (1_i), and x_{1_i} is the mole fraction of 1_i in the azeotrope (1_i, 2, 3).

2. The previously derived equations determining the composition of a positive azeotrope were adopted to a series of such azeotropes.

3. Equation (7) was essentially confirmed by comparing the only available data for the systems of benzene and methanol (isobutanol, isopropanol) with the series of aliphatic hydrocarbons.

RAW ORGANIC MATERIALS DEPARTMENT OF THE INSTITUTE OF PHYSICAL CHEMISTRY, POLISH ACADEMY OF SCIENCES

REFERENCES

- [1] W. Malesiński, *Series of binary positive azeotropes*, Bull. Acad. Polon. Sci. Cl. III, **3** (1955), 601.
- [2] W. Świętosławski, W. Malesiński, *Reduced equation characterising certain families of binary azeotropic systems*, Bull. Acad. Polon. Sci. Cl. III, **4** (1956), 159.
- [3] A. Orszagh, *Roczniki Chem.* **29** (1955), 636.
- [4] H. Skolnik, *Ind. Eng. Chem.* **40** (1948), 442.
- [5] W. Malesiński, *Composition of ternary positive azeotropes. I. II.* Bull. Acad. Polon. Sci. Cl. III, **4** (1956), 365, 523.
- [6] — *Boiling temperature of ternary positive azeotropes. I. II.* Bull. Acad. Polon. Sci. Cl. III, **4** (1956), 371, 511.
- [7] — *On the relation between the azeotropic ranges and binary regular solution constants in three-component systems*, Bull. Acad. Polon. Sci. Cl. III, **4** (1956), 303.
- [8] W. Świętosławski, A. Orszagh, *Roczniki Chem.* **26** (1952), 625.
- [9] A. Orszagh, *Roczniki Chem.* **29** (1955), 623.

Composition of Positive, Saddle and Negative Ternary Homoazeotropes

by

W. MALESIŃSKI

Presented by W. ŚWIĘTOSŁAWSKI on August 18, 1956

1. Introductory remarks. General considerations on ternary azeotropic systems forming regular solutions were advanced first by Haase [1]. He related the azeotropic composition to the vapour pressures of pure components and to the values of regular solution constants A_{12} , A_{23} and took into account the possibility and the conditions of formation of positive and positive-negative (saddle) azeotropes under isothermic conditions. In our previous papers [2], [3], the relation between, on the one hand, the composition and isobaric boiling temperature of the ternary azeotrope, and on the other, those of the binary azeotropes, taking part in the ternary system, was derived by the aid of formulae determining the activity coefficients of components forming regular solutions. We have assumed that the binary constants are related in the following way:

$$(1) \quad A_{12} + A_{23} = A_{13}.$$

This relation, derived on the basis of Świątosławski's scheme of formation of positive azeotropes [4], was tested for many pairs of binary azeotropes of ternary systems [5]. It should be emphasized that the relation may be applied only to systems in which there exist no specific interactions between unlike molecules. Therefore our considerations were limited to the systems of "ideal" positive azeotropes, in which there was not to be expected any hidden tendency to negative deviations from Raoult's law. Since we are now extending our consideration to ternary positive as well as to saddle and negative azeotropes, the specific molecular interactions cannot be overlooked, so that relation (1) becomes inapplicable.

2. Binary negative azeotropes. If one assumes that the solutions formed by components of a negative azeotrope are regular, the equations are identical with those of individual positive azeotropes [6], [7], and their series [8]. In addition, it should be assumed, as for positive systems, that the value of A_{ij} remains unchanged in successive systems of a given series. For negative systems, A_{ij} is negative. Therefore, the half-value of a symmetrical azeotropic range z_{ij} [4], which is equal to [6]:

$$(2) \quad z_{ij} = \frac{A_{ij}}{4S}$$

is also negative. The symmetry of the ranges arises from the assumption that the vaporization entropies of the components are equal to one another [9]. As previously, this simplification also is applied.

In accordance with the works cited [6], [7], the boiling temperature of both the positive and the negative azeotrope T_{Az}^{ij} is related to the azeotropic composition by

$$(3) \quad T_i - T_{Az}^{ij} = z_{ij} x_j^2,$$

where T_i is the boiling temperature of pure component i and x_j — the mole fraction of j in the azeotrope. The composition of both positive and negative azeotropes depends on the differences in boiling temperatures of pure components in the following way;:

$$(4) \quad x_i = \frac{1}{2} + \frac{T_j - T_i}{2z_{ij}}.$$

The negative azeotropic depressions are correlated in accordance with (3) by the equation:

$$(5) \quad \sqrt{|T_i - T_{Az}^{ij}|} + \sqrt{|T_j - T_{Az}^{ij}|} = \sqrt{|z_{ij}|}.$$

In that way, the value of the range may be easily estimated, even for solutions which cannot be considered as regular; that means that the parabolic relation (3) is, for them, not fulfilled.

This results from the fact that equation (5) is valid independently of which concentration units are used to express the azeotropic composition.

3. Relation between the composition of ternary and binary azeotropes.

If the vapour obeys perfect gas laws, the following relation holds at the azeotropic point:

$$(6) \quad p_i^0(T_{Az}) f_i(T_{Az}, x_{Az}) = P.$$

p_i^0 is the vapour pressure of i , f_i — the activity coefficient of this component in the azeotropic mixture at the boiling temperature T_{Az} under

the pressure P . The activity coefficient of i in the regular solution composed of i, j, k , is:

$$(7) \quad RT \ln f_i = A_{ij}x_j^2 + A_{ik}x_k^2 + (A_{ij} + A_{ik} - A_{jk})x_jx_k.$$

Using equation (6), and introducing the parameters z , we obtain in the same way as in (2) the equation:

$$(8) \quad T_{Az}^{ij} - T_i + z_{ij}x_j^2 + z_{ik}x_k^2 + (z_{ik} + z_{ij} - z_{jk})x_jx_k = 0.$$

On deriving this, we assume that the vaporization entropy remains constant in the temperature interval T_i to T_{Az} . The analogous equation for component j is:

$$(9) \quad T_{Az}^{(ij)} - T_j + z_{ij}x_i^2 + z_{jk}x_k^2 + (z_{jk} + z_{ij} - z_{ik})x_jx_k = 0.$$

Subtracting (9) from (8) we obtain:

$$(10) \quad T_j - T_i + z_{ij} - 2z_{ij}x_i + x_k(z_{ik} - z_{ij} - z_{jk}) = 0.$$

Dividing by $2z_{ij}$, $z_{ij} \neq 0$, we transform it as follows:

$$(11) \quad \frac{T_j - T_i}{2z_{ij}} + \frac{1}{2} - x_i + x_k \frac{z_{ik} - z_{ij} - z_{jk}}{2z_{ij}} = 0.$$

Taking into account the relation (4), valid in the case of positive as well as negative azeotropy, (11) may be rewritten as follows:

$$(12) \quad x_i^{(ij)} - x_i + x_k \frac{z_{ik} - z_{ij} - z_{jk}}{2z_{ij}} = 0.$$

The analogous equation correlating the composition of the ternary azeotrope with the mole fraction of component k in the binary azeotrope (j, k) has the form:

$$(13) \quad x_k^{(j,k)} - x_k + x_i \frac{z_{jk} - z_{jk} - z_{ij}}{2z_{jk}} = 0.$$

Solving the system of equations (12) and (13), we obtain:

$$(14) \quad x_i = \frac{x_i^{(i,j)} + x_k^{(j,k)} \frac{z_{ik} - z_{ij} - z_{jk}}{2z_{ij}}}{1 - \frac{(z_{ik} - z_{ij} - z_{jk})^2}{4z_{ij}z_{jk}}}.$$

The general formula (14) relates the composition of the ternary azeotrope to the compositions of the binary ones. The azeotropes may be positive, saddle or negative.

4. General analysis of the results. Substituting the subscripts i, j, k by 1, 2, and 3, we get, in accordance with (12), a set of equations correlating the respective mole fractions in binary and ternary mixtures:

$$(15a) \quad x_1^{(1,2)} - x_1 + x_3 a = 0, \quad (15b) \quad x_3^{(2,3)} - x_3 + x_1 b = 0,$$

$$(15c) \quad x_1^{(1,3)} - x_1 + x_2 c = 0, \quad (15d) \quad x_2^{(2,3)} - x_2 + x_1 d = 0,$$

$$(15e) \quad x_2^{(1,2)} - x_2 + x_3 e = 0, \quad (15f) \quad x_3^{(1,3)} - x_3 + x_2 f = 0,$$

in which the parameters a, b, c, d, e, f , are:

$$(16a) \quad a = \frac{z_{13} - z_{12} - z_{23}}{2z_{12}}, \quad (16b) \quad b = \frac{z_{13} - z_{12} - z_{23}}{2z_{23}},$$

$$(16c) \quad c = \frac{z_{12} - z_{23} - z_{13}}{2z_{13}}, \quad (16d) \quad d = \frac{z_{12} - z_{23} - z_{13}}{2z_{23}},$$

$$(16e) \quad e = \frac{z_{23} - z_{12} - z_{13}}{2z_{12}}, \quad (16f) \quad f = \frac{z_{23} - z_{12} - z_{13}}{2z_{13}},$$

and the conditions are fulfilled:

$$(17a) \quad a + e + 1 = 0, \quad (17b) \quad b + d + 1 = 0, \quad (17c) \quad c + f + 1 = 0.$$

One can see that the constants a, b, c as well as d, e , and f are connected one with another. The relation between a, b and c is expressed as follows:

$$(18) \quad 2abc + ac + bc + ab + a = 0,$$

resulting from (16a), (16b) and (16c).

Solving the successive pairs of equations (15), and substituting the symbols i, j, k by 1, 2 and 3, we get the three pairs of relations:

$$(19a) \quad x_1 = \frac{x_1^{(1,2)} + ax_3^{(2,3)}}{1 - ab} \quad x_3 = \frac{x_3^{(2,3)} + dx_1^{(1,2)}}{1 - ab},$$

$$(19b) \quad x_1 = \frac{x_1^{(1,3)} + cx_2^{(2,3)}}{1 - cd} \quad x_2 = \frac{x_2^{(2,3)} + dx_1^{(1,3)}}{1 - cd},$$

$$(19c) \quad x_2 = \frac{x_2^{(1,2)} + ex_3^{(1,3)}}{1 - ef} \quad x_3 = \frac{x_3^{(1,3)} + fx_2^{(1,2)}}{1 - ef}.$$

It is clear, that if the system has all the properties of regular solutions, the three pairs of relations are equivalent. However, in the case of deviations from such properties, the three relations must lead to dif-

ferent results, depending on which pair of binary azeotropes is chosen for the calculations. The fact, that it is sufficient, in order to find the composition of ternary azeotrope to know the composition of two binary azeotropes; arises from the following condition:

$$(20) \quad 2z_{12}x_2^{(1,2)} + 2z_{23}x_3^{(2,3)} - 2z_{13}x_3^{(1,3)} + (z_{13} - z_{23} - z_{13}) = 0.$$

This is derived by substituting the respective temperature differences in (4) into the identity: $(T_1 - T_2) + (T_2 - T_3) - (T_1 - T_3) = 0$.

Attention should be paid to the fact that for the formation of the ternary azeotrope the existence of all three binary azeotropes is not necessary. One and even two systems may be zeotropic. In this case, when applying equations (19), unreal azeotropic concentrations should be ascribed to zeotropic systems (mole fractions $x < 0$ or $x > 1$). The value of this "unreal concentration" may be estimated by means of (4). Then, if necessary, the azeotropic range may be determined from the data available for another system of the same series, if specific interactions between unlike molecules are absent, by means of 1.

Constants a, b, c, d, e, f in (16), characterizing the properties of the given system, may vary within wide limits, and may be positive or negative. In particular, if the value of one of the ranges is close to zero, and the difference of the remaining two differs significantly from zero, the value of the respective constant is very large, reaching $\pm \infty$ if $z \rightarrow 0$. The system with $z = 0$ is an ideal or pseudoideal system.

5. "Ideal" positive azeotropes. The previously considered positive azeotropic systems [2], [3], were characterized by the absence of specific interactions between unlike molecules. For such systems relation (1) is valid and also the following relation resulting from by means of (2):

$$(21) \quad z_{12} + z_{23} = z_{13}.$$

Azeotropic systems fulfilling relation (21) may be termed "ideal" positive azeotropic systems. It will be shown, further, in what way the general relations for composition of ternary azeotropes are simplified in the case of "ideal" positive azeotropes. If relation (21) is fulfilled, the respective constants are:

$$a = 0, \quad b = 0 \quad c = -\frac{z_{23}}{z_{13}}, \quad d = -1, \quad e = -1, \quad f = -\frac{z_{12}}{z_{13}}.$$

Then, according to (19a), the composition of the ternary azeotrope is,

$$(22) \quad x_1 = x_1^{(1,2)} \quad x_3 = x_3^{(2,3)}.$$

Component 2 is the main component, i. e., it is characterized by the medium value of internal pressure. A more detailed discussion of (22),

and the extension of it to systems with different vaporization entropies of the components, has been given elsewhere [2], [3].

If a series of azeotropes is taken into consideration [10], [11], i. e., if one of the components is replaced by its successive homologues, and if the parameters a and b (and therefore even c) remain constant in the series of systems, then the ternary azeotropic points of this series should lie on one straight line, determined by equations (15). This was found to agree with experimental data recorded by Orszagh [12], in the case of series of positive azeotropes formed by benzene, alcohols (metanol, isopropanol, isobutanol) and a gasoline fraction. The results have been previously discussed [11]. Our predictions are also approximately in conformity with experimental data obtained by Zięborak [13] for a series of saddle azeotropes formed by acetic acid, pyridine and the series of normal paraffins. A more detailed discussion of other series of ternary azeotropes has been prepared for publication.

Summary

1. On the assumption that the activity coefficients of the components of a ternary system are determined by formulae valid for regular solutions, an equation for the composition of ternary azeotropes was derived.

2. It is shown, that under isobaric conditions the mole fraction x_i of component i in the ternary azeotropic mixture is related to $x_i^{(i,j)}$ and $x_k^{(j,k)}$ in the respective binary azeotropic mixtures ij and jk , as follows:

$$(14) \quad x_i = \frac{x_i^{(i,j)} + x_k^{(j,k)} \frac{z_{ik} - z_{ij} - z_{jk}}{2z_{ij}}}{1 - \frac{(z_{ik} - z_{jk} - z_{ij})^2}{4z_{ij}z_{jk}}},$$

z_{ij} , z_{ik} , z_{jk} being the half-values of the azeotropic ranges in systems ij , ik and jk

RAW ORGANIC MATERIALS DEPARTMENT OF THE INSTITUTE OF PHYSICAL CHEMISTRY, POLISH ACADEMY OF SCIENCES

REFERENCES

- [1] R. Haase, *Z. physik. Chem.* **195** (1950), 362.
- [2] W. Malesiński, *Composition of ternary positive azeotropes*, I, II. *Bull. Acad. Polon. Sci. Cl. III*, **4** (1956), 365, 523.
- [3] — *Boiling temperature of positive azeotropes*, I, II. *Bull. Acad. Polon. Sci. Cl. III*, **4** (1956), 371, 609.
- [4] W. Świątosławski, *Bull. Soc. Chim. Belges*, **62** (1953), 10.
- [5] W. Malesiński, *On the relation between the azeotropic ranges and binary regular solution constants in three-component systems*, *Bull. Acad. Polon. Sci. Cl. III*, **4** (1956), 303.
- [6] — *Series of binary positive azeotropes*, *Bull. Acad. Polon. Sci. Cl. III*, **3** (1955), 601.

- [7] W. Świątosławski, W. Malesiński, *Reduced equation characterising certain families of binary azeotropic systems*, Bull. Acad. Polon. Sci. Cl. III, **4** (1956), 159.
- [8] — Roczniki Chem. **26** (1952), 605.
- [9] W. Malesiński, *Series of binary azeotropic systems characterised by symmetrical azeotropic ranges*, Bull. Acad. Polon. Sci. Cl. III, **4** (1956), 295.
- [10] W. Świątosławski, *Azeotropia i Poliazeotropia*; in press.
- [11] — W. Malesiński, *Azeotropic depression as a function of the composition of series of ternary positive azeotropes*, Bull. Acad. Polon. Sci. Cl. III.
- [12] A. Orszagh, Roczniki Chem. **29** (1955), 623.
- [13] K. Zięborek, *Azeotropic and polyazeotropic systems. XXI. A series of saddle azeotropes formed by acetic acid, pyridine and paraffinic hydrocarbons*, Bull. Acad. Polon. Sci. Cl. III, **3** (1955), 531.

Boiling Temperatures of Positive, Saddle and Negative Ternary Homoazeotropes

by

W. MALESIŃSKI

Presented by W. ŚWIEȚOSŁAWSKI on August 18, 1956

1. Introductory remarks. Equations are here derived for determining the boiling temperature of ternary azeotropic systems in which the activity coefficients of the components are fixed by formulae valid for regular solutions. Our considerations refer to positive, positive-negative (saddle) and negative azeotropes. It will be remembered that previously [1] we obtained the equation for the boiling temperature of "ideal" positive azeotropes, i. e., those in which there existed no specific interactions between unlike molecules [2], [3]. The equation has the form:

$$(1) \quad T_2 - T_{Az}^{(123)} = (T_2 - T_{Az}^{(1,2)}) + (T_2 - T_{Az}^{(2,3)}).$$

T_2 is the boiling temperature of the main component [2], [4], $T_{Az}^{(1,2,3)}$ — that of the ternary and $T_{Az}^{(1,2)}$ and $T_{Az}^{(2,3)}$ — those of the binary azeotropes containing component 2.

The purpose of the present paper is to derive a more general relation determining the boiling temperature not only of an ideal positive ternary azeotrope but of all azeotropes, positive, saddle and negative, treated as regular solutions.

2. General equations. Our starting point is the equation relating the boiling temperature of the ternary azeotrope $T_{Az}^{(1,2,3)}$ to the mole fractions x_i and x_k of the components i and k in the azeotrope (i, j, k) [3]:

$$(2) \quad T_{Az}^{(123)} - T_j + \frac{A_{ij}}{\Delta S} x_i^2 + \frac{A_{jk}}{\Delta S} x_k^2 + \left(\frac{A_{ij}}{\Delta S} + \frac{A_{jk}}{\Delta S} - \frac{A_{ik}}{\Delta S} \right) x_i x_k = 0,$$

in which A_{ij} , A_{jk} and A_{ik} are the regular solution constants, ΔS — the vaporisation entropy of the pure components at their boiling temperature, as assumed, equal for all of them.

Introducing the azeotropic range z as a parameter, we may rewrite equation (2) as follows:

$$(3) \quad T_j - T_{\text{Az}}^{(1,2,3)} = z_{ij}x_i^2 + z_{jk}x_k^2 + (z_{ij} + z_{jk} - z_{ik})x_i x_k.$$

Rearranging it, we obtain:

$$(4) \quad T_j - T_{\text{Az}}^{(1,2,3)} = z_{ij}x_i \left(x_i - x_k \frac{z_{ik} - z_{ij} - z_{jk}}{2z_j} \right) + z_{jk}x_k \left(x_k - x_i \frac{z_{ik} - z_{ij} - z_{jk}}{2z_{jk}} \right).$$

Since the compositions of the binary and ternary azeotropes are correlated by [3]:

$$(5a) \quad x_i^{(i,j)} - x_i + x_k \frac{z_{ik} - z_{ij} - z_{jk}}{2z_{ij}} = 0,$$

$$(5b) \quad x_k^{(k,j)} - x_k + x_i \frac{z_{ik} - z_{kj} - z_{jk}}{2z_{kj}} = 0,$$

equation (4) becomes:

$$(6) \quad T_j - T_{\text{Az}}^{(1,2,3)} = z_{ij}x_i x_i^{(i,j)} + z_{jk}x_k x_k^{(k,j)},$$

where $x_i^{(i,j)}$ and $x_k^{(k,j)}$ are the mole fractions of components i and k in azeotropes (i,j) and (k,j) .

Taking into account the relations for binary azeotropes [5]:

$$(7) \quad T_j - T_{\text{Az}}^{(i,j)} = z_{ij}[x_i^{(i,j)}]^2; \quad T_j - T_{\text{Az}}^{(j,k)} = z_{jk}[x_k^{(j,k)}]^2,$$

we obtain

$$(8) \quad T_j - T_{\text{Az}}^{(i,j,k)} = (T_j - T_{\text{Az}}^{(i,j)}) \frac{x_i}{x_i^{(ij)}} + (T_j - T_{\text{Az}}^{(j,k)}) \frac{x_k}{x_k^{(kj)}}.$$

For the case in which "ideal" positive azeotropes are formed, relation (8) transforms itself into (1) because for such azeotropes [1], [3],

$$x_1^{(12)} = x_1 \quad \text{and} \quad x_3^{(23)} = x_3.$$

For correlating the boiling temperature of the ternary azeotrope with the parameters exclusively characterising the properties of binary systems, the previously derived formulae [3] must be introduced:

$$(9a) \quad x_i = \frac{x_i^{(ij)} + x_k^{(jk)} \frac{z_{ik} - z_{ij} - z_{jk}}{2z_{ij}}}{1 - \frac{(z_{ik} - z_{ij} - z_{jk})^2}{4z_{ij}z_{jk}}},$$

$$(9b) \quad x_k = \frac{x_k^{(jk)} + x_i^{(ij)} \frac{z_{ik} - z_{ij} - z_{jk}}{2z_{jk}}}{1 - \frac{(z_{ik} - z_{ij} - z_{jk})^2}{4z_{ij}z_{jk}}}.$$

Substituting in (6) the expressions for the composition x_i and x_k of a ternary azeotrope, we obtain:

$$(10) \quad T_j - T_{Az}^{(i,j,k)} = \frac{z_{ij} x_i^{(i,j)^2} + x_i^{(i,j)} x_k^{(j,k)} (z_{ik} - z_{ij} - z_{jk}) + z_{jk} x_k^{(k,j)^2}}{1 - \frac{(z_{ij} - z_{ij} - z_{jk})^2}{4z_{ij}z_{jk}}}$$

Replacing the mole fractions $x_i^{(i,j)}$ and $x_k^{(j,k)}$ by the expressions obtained from (7), we get the final relation:

$$(11) \quad T_j - T_{Az}^{(i,j,k)} = \frac{\delta_{ji} + \sqrt{\frac{\delta_{ji}}{z_{ij}}} \sqrt{\frac{\delta_{jk}}{z_{jk}}} (z_{ik} - z_{ij} - z_{jk}) + \delta_{jk}}{1 - \frac{(z_{ik} - z_{ij} - z_{jk})^2}{4z_{ij}z_{jk}}},$$

in which $\delta_{ji} = T_j - T_{Az}^{(ij)}$ and $\delta_{jk} = T_j - T_{Az}^{(jk)}$ are the azeotropic depressions in binary systems (ij) and (jk) . Since the azeotropic range z_{ij} is determined by

$$(12) \quad \sqrt{|\delta_{ij}|} + \sqrt{|\delta_{ji}|} = \sqrt{|\delta_{ij}|},$$

where $\delta_{ij} = T_i - T_{Az}^{(i,j)}$ and $\delta_{ji} = T_j - T_{Az}^{(i,j)}$, equation (11) makes it possible to calculate the boiling temperature of ternary azeotropes, knowing only the boiling temperatures of pure components and their binary azeotropes.

3. Comparison with experimental results. Positive azeotropes. The comparison of data, calculated and experimental, for 42 ternary azeotropes [1], quoted in Horsley's Tables [6], confirmed in general the usefulness of relation (1), although several systems showed certain deviations from the properties of "ideal" positive azeotropes. In such cases, the value $\Delta z = z_{13} - z_{12} - z_{23}$ differed from zero.

Fig. 1 represents the experimental values $\delta_{213} = T_2 - T_{Az}^{(123)}$ plotted against the calculated values δ_{213} , according to (1). Now, the same comparison is carried out with the δ_{213} - values calculated by means of equation (11) (Fig. 2). The latter is more general. It takes into consideration the specific interactions between unlike molecules, and may be rewritten as follows:

$$(13) \quad T_2 - T_{Az}^{(123)} = \frac{\delta_{21} + \sqrt{\frac{\delta_{21}}{z_{12}}} \sqrt{\frac{\delta_{23}}{z_{23}}} \Delta z + \delta_{23}}{1 - \frac{\Delta z^2}{4z_{12}z_{23}}}.$$

The symbols may be ascribed directly to any of the components, without leading to inconsistent numerical results even if there exists a difference between the real and the calculated boiling temperature of the ternary azeotrope.

Table I contains the experimental values of boiling temperatures of 42 ternary azeotropes, t_{Az} (exp.), values calculated, t_{Az} (calc) from (13), and the differences between them, Δt . For convenience, values $\Delta t'$, calculated by means of equation (1), are added to the table. The diagram of t_{Az} (exp), vs. t_{Az} (calc), according to (13), is shown in Fig. 2.

TABLE I

No.	$t_{Az}^{(1,2,3)}$ exp.	$t_{Az}^{(1,2,3)}$ calc.	Δt	$\Delta t'$	No.	$t_{Az}^{(1,2,3)}$ exp.	$t_{Az}^{(1,2,3)}$ calc.	Δt	$\Delta t'$
1	35.95	36.6	+ 0.65	+0.75	22	127.5	126.99	— 0.51	—0.35
2	37.2	37.25	+ 0.05	+1.15	23	24.1	20.12	— 3.98	+2.35
3	33.92	34.45	+ 0.53	+0.53	24	16.95	16.19	— 0.76	—1.8
4	37.0	37.51	+ 0.51	—0.45	26	24	24.31	+ 0.31	+0.15
5	37.0	36.97	— 0.03	—0.95	27	24	10.53	—13.47	+3.8
6	24.72	15.7	— 9.0	+4.6	28	20.4	—19.71	—40.11	+6.6
7	35.55	36.60	+ 1.14	+1.27	29	64.3	64.85	+ 0.55	+3.25
8	24.0	23.90	— 0.1	—0.35	30	162.45	163.75	+ 1.3	+5.1
9	21.5	21.84	+ 0.34	—0.15	31	165.5	165.09	— 0.41	—0.9
10	38.2	39.28	+ 1.08	+3.5	32	68.3	68.55	+ 0.25	+1.45
11	35.2	34.88	— 0.32	—0.65	33	74	73.88	— 0.12	+0.32
12	58.3	58.33	+ 0.03	0.0	34	154.5	153.24	— 1.26	—3.25
13	38.5	38.74	+ 0.24	+0.52	35	152.3	152.41	+ 0.11	+0.3
14	99	97.65	— 1.35	—1.35	36	150.4	150.42	+ 0.02	0.0
15	31.4	31.43	+ 0.03	+0.2	37	146.4	146.75	+ 0.35	—0.5
16	53.9	54.3	+ 0.4	+0.35	38	143.9	146.47	+ 2.57	+1.7
17	52.0	52.52	+ 0.52	+0.3	39	168.7	168.78	+ 0.08	—2.2
18	51.1	51.90	+ 0.80	+1.15	40	168.7	168.76	+ 0.06	—2.15
19	23	11.59	—11.41	+0.45	41	197	194.64	— 2.46	+2.85
20	116.0	115.86	+ 0.14	+1.85	42	172.5	172.53	+ 0.03	+0.95
21	117.6	118.56	+ 0.96	+0.15					

A better agreement is obtained by employing equation (13) instead of (1). The former fits the data of 29 systems within $\pm 1^\circ\text{C}$., with a mean absolute deviation equal to 0.32°C . The latter fits, within this limit, 22 systems with a mean deviation 0.42°C .

The mean deviations lie within the limits of experimental errors. In fact, the boiling temperatures of pure components and binary azeotropes are, generally speaking, determined with a degree of error amounting to about 0.05 to 0.1°C . Then, according to (1), the boiling temperature of the ternary azeotrope has a degree of error within 0.2 to 0.4°C . In the case, however, of systems 6, 19, 23, 27, 28 and 41 the discrepancies are very large and, further, they are much greater when using equation (13). This strange phenomenon may be explained as follows.

To begin with, as shown in Table II, the determined boiling temperatures of the ternary azeotropes lie close to those of the lowest boiling temperatures of the binary azeotropes (1,3). These were found by Lecat [7]

by fractional distillation. Since the boiling temperature differences are very small, the separating power of the column used must be very high.

TABLE II

No. of syst.	6	19	23	27	28	41
$t_{Az}^{(1,3)}$	24.75	22.5	24.3	24.3	21.8	198.7
$t_{Az}^{(1,2,3)}$	24.7?	23?	24.1	24	20.4	197?

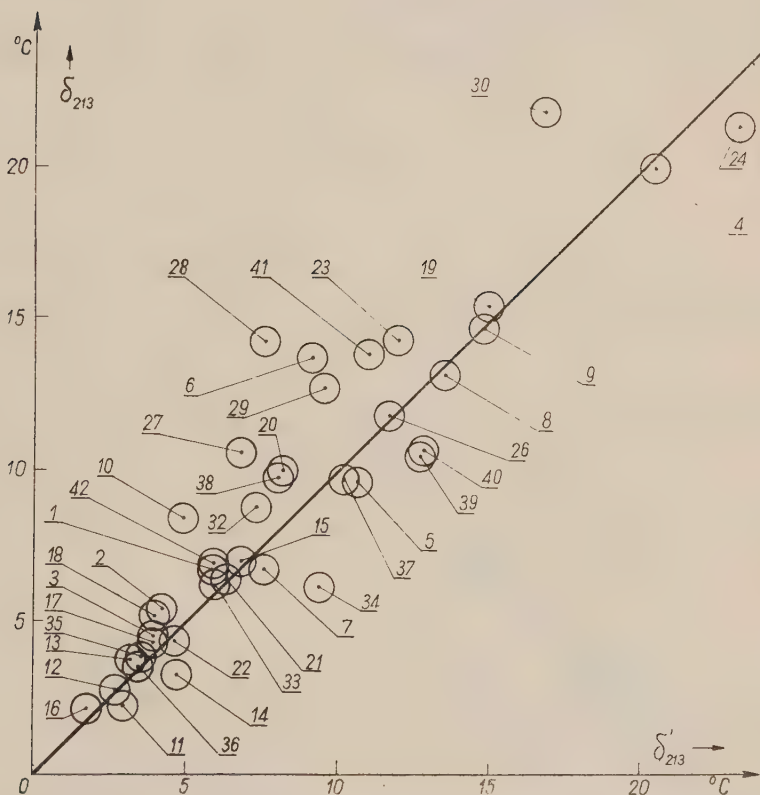


Fig. 1. Experimental values of azeotropic depressions δ_{213} as a function of δ'_{213} calculated from the relation (1)

We believe that the appearance of component 2 in the receiver is due only to insufficient fractionation [8] and that the azeotropes collected in Table II do not, in fact, exist. Moreover the absence of systems deviating to an intermediate extent supports the above conclusion.

4. Positive-negative (saddle) azeotropes. For the purpose of verification, data given by Zięborak [9] and Kurtyka [10] for four saddle azeotropes were used. In Table III, are collected the boiling temperatures

TABLE III

No.	system	t_1	t_2	t_3	$t_{Az}^{(1,2)}$	$t_{Az}^{(2,3)}$	$t_{Az}^{(1,3)}$	$t_{Az}^{exp.}$	$t_{Az}^{calc.}$	Δt
1.	<i>n</i> -Heptane-pyridine-acetic acid	98.4	115.4	118.2	95.5	138.1	91.8	96.5	95.6	-0.9
2.	<i>n</i> -octane-pyridine-acetic acid	125.6	115.4	118.2	109.5	138.1	105.8	117.7	113.2	-2.5
3.	<i>n</i> -nonane-pyridine-acetic acid	150.5	115.4	118.2	115.1	138.1	113.0	128.0	124.6	-3.4
4.	2, 4, 6-collidine- <i>o</i> -cresol-glycol	171.30	191.00	197.45	197.2	189.52	170.5	189.65	189.85	+0.20



Fig. 2. Experimental values of azeotropic depressions as a function of δ'_{213} calculated from the relation (13)

of pure components t_1 , t_2 and t_3 , those of the binary azeotropes $t_{Az}^{(1,2)}$, $t_{Az}^{(2,3)}$ and $t_{Az}^{(1,3)}$, and those of the saddle azeotropes: experimental t_{Az}^{exp} and calculated t_{Az}^{calc} according to (13). The final column contains the difference of these values.

Considerable differences observed in systems 2 and 3 are probably due to almost complete association of acetic acid in both, and in the liquid and the vapour phase. The smaller difference in the case of system 1 is due to low acetic acid content in the azeotropic mixture. It is of interest that, for system 4, composed of three polar substances, good agreement was obtained.

Summary

1. Assuming as hitherto [1] that the activity coefficients of components i, j, k , of a ternary system are determined by the equations for regular solutions, a new equation was derived for the boiling temperature of ternary azeotropes T_{Az}^{ijk} . Under isobaric conditions, T_{Az}^{ijk} is related to the azeotropic depressions of binary systems by

$$(11) \quad T_j - T_{Az}^{ijk} = \frac{\delta_{ij} + \sqrt{\frac{\delta_{ji}}{z_{ij}}} \sqrt{\frac{\delta_{jk}}{z_{jk}}} (z_{ik} - z_{ij} - z_{jk}) + \delta_{jk}}{1 - \frac{(z_{ik} - z_{ij} - z_{jk})^2}{4z_{ij}z_{jk}}},$$

z_{ij} , z_{jk} , and z_{ik} being the half-values of the symmetric azeotropic ranges.

2. Equation (11) was tested by comparison with the data for 41 ternary systems compiled by Horsley. In most cases, a good agreement was found. It is probable that 6 ternary azeotropes cited by Horsley do not exist.

3. The results calculated by using equation (12) were compared with those found for 4 positive-negative (saddle) azeotropes. When acetic acid was one of the components, a poor agreement with experimental data was observed.

RAW ORGANIC MATERIALS DEPARTMENT OF THE INSTITUTE OF PHYSICAL CHEMISTRY, POLISH ACADEMY OF SCIENCES

REFERENCES

- [1] W. Malesiński, *Boiling temperature of positive azeotropes*, Bull. Acad. Polon. Sci. Cl. III, **4** (1956), 371.
- [2] — *On the relation between the azeotropic ranges and binary regular solution constants in three-component systems*, Bull. Acad. Polon. Sci. Cl. III, **4** (1956), 303.
- [3] — *Composition of ternary positive, positive-negative and negative azeotropes*, Bull. Acad. Polon. Sci. Cl. III, 703.
- [4] W. Świątosławski, Bull. Soc. Chim. Belges **62** (1953), 10.
- [5] W. Malesiński, *Series of binary positive azeotropes*, Bull. Acad. Polon. Sci. Cl. III, **3** (1955), 601.
- [6] L. H. Horsley, *Azeotropic data*, Am. Chem. Soc. Washington 1952.
- [7] M. Lecat, *L'Azeotropisme*, Brussels, Lamartin 1918.
- [8] W. Świątosławski, A. Orszagh, Roczn. Chem. **26** (1952), 608.
- [9] K. Zięborak, *Azeotropic and polyazeotropic systems XXI*, Bull. Acad. Polon. Sci. Cl. III, **3** (1955), 531.
- [10] Z. Kurtyka, *The positive-negative azeotrope: Glycol-o-Cresol 2, 4, 6-Colldine XXIII*, Bull. Acad. Polon. Sci. Cl. III, **4** (1956), 49.

Kinetics and Mechanism of Cementation Processes. I. Model Lead-Zinc Cementation Cell

by

J. SĘDZIMIR and M. PAWEŁKOWA

Presented by M. ŚMIAŁOWSKI on July 4, 1956

1. The kinetics and mechanism of precipitation of metals by metals from aqueous solutions have been the subject of research work for more than twenty years. The last ten years or so have brought tentatives to present a general theory of cementation processes.

Plaksin and Suvorovskaya [1] claim that cementation processes, without evolution of hydrogen, proceed in agreement with the kinetic equation for first order reactions. Rotinyan and Drozdova [2], [3] believe cementation products slow down the process. Bulakh and Dratchevskaya [4] have concluded that the reaction products accelerate the process. Still other views have been put forward by Piontelli [5], Glicksman, Mouquin, King [6] and by Krzyżański [7].

In researches hitherto undertaken the precipitating metal was immersed in a solution of salts of the metal precipitated, and changes in the concentration of the latter were determined analytically in the course of cementation.

Discordant views on kinetics and mechanism of cementation as well as some vagueness in our own work [8], [9], [10] have prompted us to investigate the process in a model system. Expediency of investigating kinetics of cementation in a Daniell-type cell was also mentioned by Piontelli in one of his recent papers [5].

The difficulties met with when investigating the kinetics in this system consist in that the cemented metal is deposited simultaneously on both electrodes.

2. **Method of measurements.** One of the earlier publications [10] dealt with precipitation with zinc of lead from a chloride solution (PbCl_2 in 20 per cent NaCl). This work was undertaken with a view to investigating the system mentioned in a model cell.

In order to prevent the metal precipitated from depositing simultaneously on both the cathode and the anode, the electrodes were separated by a porous diaphragm. Zinc and lead sheets served as anode and cathode respectively. We used 20 per cent solution of NaCl as anolyte and 20 per cent NaCl + PbCl₂ as catholyte. The electrodes were connected over a miliamperemeter of low internal resistance (0.5 ohm). In both electrode compartments the solution was stirred at a constant rate (100 revolutions per minute), and the entire system was kept in a thermostat at a temperature of $25 \pm 0.5^\circ\text{C}$. While the process was in progress the current in the system and the potentials of both electrodes (against a standard calomel electrode) were measured. A simplified diagram of the arrangement is presented in Fig. 1.

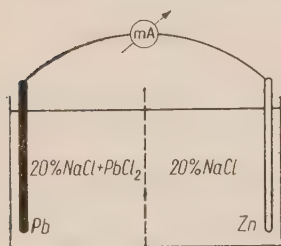


Fig. 1. Model lead-zinc cementation cell

In the course of the process, lead evolved on the cathode in the form of a notably comminuted sponge. Readings were taken at various concentrations of Pb in the catholyte and at different surfaces of the electrodes.

3. Results. Results obtained during precipitation from solutions containing various initial amounts of lead (1-6 g. Pb/l.) are presented in Figs. 2 and 3.

It should be mentioned that while current readings were taken, variations within ± 10 mA. were observed; only on the horizontal sections of the curves (4 and 6 g. Pb/l.) the variations were ± 2 mA.

For reasons of clarity, merely the curves relating to the initial concentrations of 6 and 1 g. Pb/l. are presented in Fig. 3. The curves for 4 g. Pb/l. and for 6 g. Pb/l. resemble each other (the horizontal section is shorter). Curves for 2.0 and 1.5 g. Pb/l. resemble that for 1 g. Pb/l., at higher concentrations, however, the maximum is shifted to more negative values.

Precipitation from concentrations low in lead content (1-2 g. Pb/l.) with the aid of lead cathodes of different surfaces (within 10-30 sq. cm.) yielded curves similar in shape but differing one from the other in the values of " i_{mx} " and in the time after which the former value was obtained (" t_{mx} "). The data are recorded in Table I.

Tentatives to present the results obtained from dilute solutions (up to 2 g. Pb/l.) in various systems of co-ordinates resulted in the following empirical equations:

1. $i_{mx} = A_1 c_0$ — the maximum current is a linear function of initial concentration; the slope of curve A_1 is the function of the initial cathode surface.

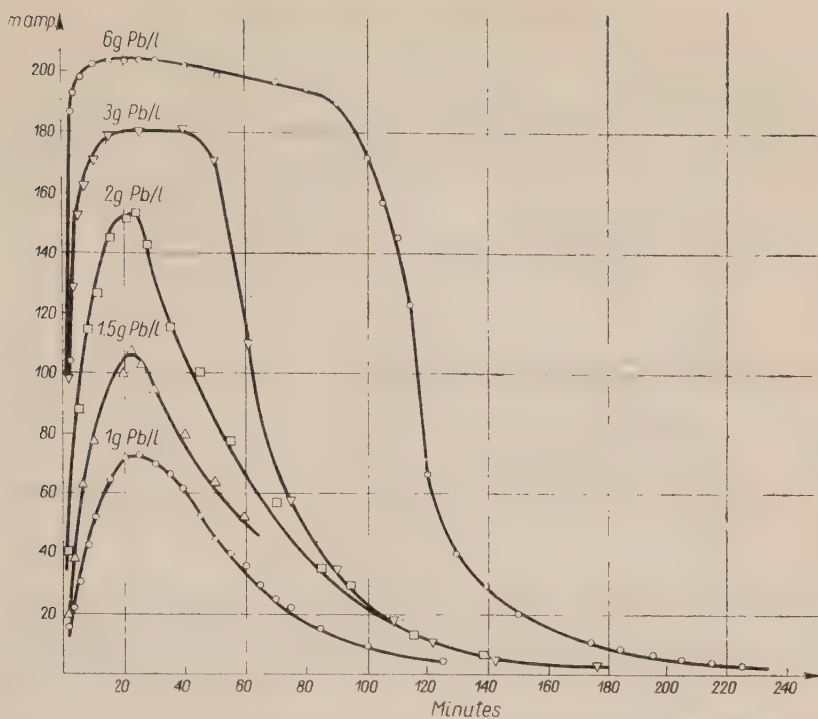


Fig. 2. Current intensity in the model cell as a function of time at different initial lead concentrations in the catholyte. Surface of the cathode = surface of the anode = 30 sq. cm.; $t = 25^\circ\text{C}$.; stirred at 100 revolutions per min.

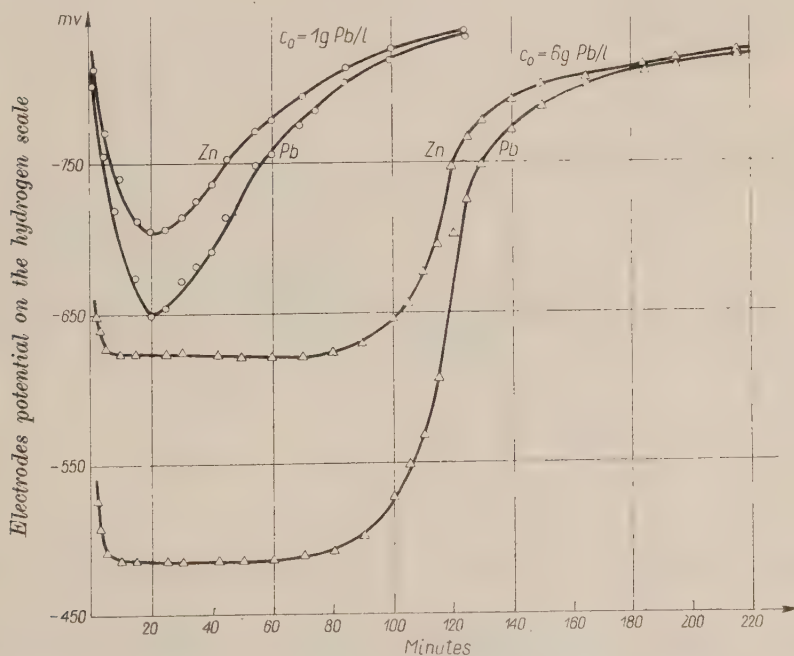


Fig. 3. Potential of the electrodes in the model cell as a function of time at different initial lead concentrations. Surface of cathode = surface of anode = 30 sq. cm.; $t = 25^\circ\text{C}$.; stirred at 100 revolutions per min.

2. $\lg i_{mx} = A_2 + B_2 F_0$ — the logarithm of the maximum current intensity is a linear function of the initial cathode surface; the slope

of curve B_2 does not depend on the initial concentration, A_2 , however, is a function of c_0 .

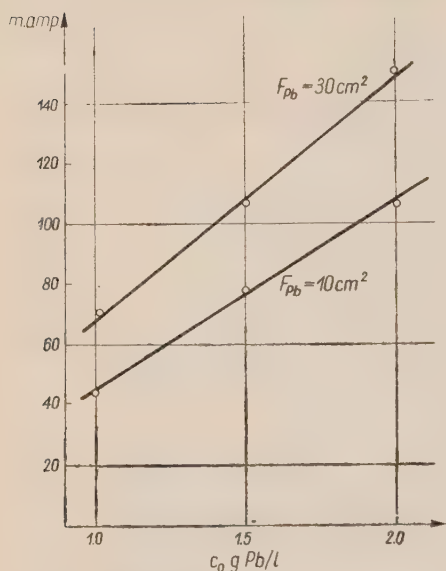


Fig. 4. Maximum current obtained in the model cell expressed as a function of the initial lead ion concentration in the catholyte. $F_{Zn} = 30 \text{ sq. cm.}$; $t = 25^\circ \text{C.}$; stirring at 100 revolutions/min.

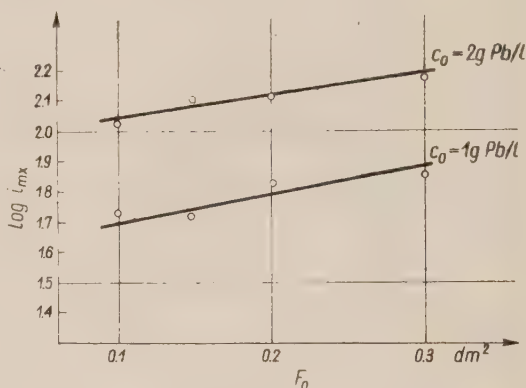


Fig. 5. The logarithm of the current maximum value obtained in the model cell expressed as a function of initial lead cathode surface. $c_0 = 1$ and 2 g. Pb/l. ; $t = 25^\circ \text{C.}$; $F_{Zn} = 30 \text{ sq. cm.}$; stirring at 100 revolutions/min.

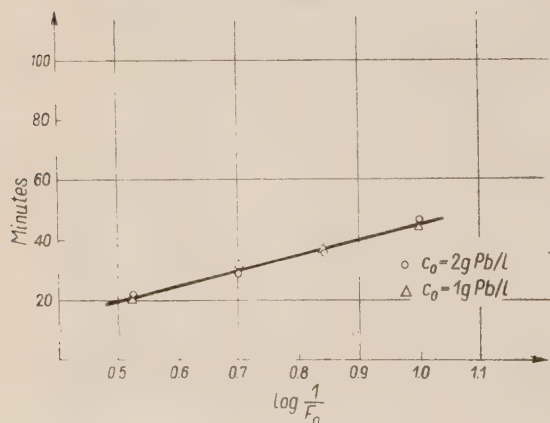


Fig. 6. The time after which maximum current is attained in the model cell expressed as a function of the logarithm of the reverse initial surface of the lead cathode. $c_0 = 1$ and 2 g. Pb/l. ; $t = 25^\circ \text{C.}$; $F_{Zn} = 30 \text{ sq. cm.}$; stirring at 100 revolutions/min.

3. $t_{mx} = A_3 + B_3 \lg \frac{1}{F_0}$ — the time corresponding to the current maximum value is a linear function of $\lg \frac{1}{F_0}$; coefficients A_3 and B_3 do not depend upon initial concentration.

The relations discussed in equations 1—3 are presented in Figs. 4—6.

TABLE I

Precipitation with the aid of cathodes of various surfaces; $t = 25^\circ\text{C}$.; $F_{Zn} = 30$ sq. cm.; stirring at 100 revolutions/min.

Initial concentration of the catholyte	F_{Pb} in sq. cm.	i_{mx} in mA	t_{mx} in min.
$c_0 = 1$ g. Pb/l	30	72	21
	20	68	32
	14.4	52	36
	10	53	44
$c_0 = 2$ g. Pb/l	30	152	23
	20	130	28
	14.4	125	35
	10	106	44

Tentative theoretic explanation of the relations and of the presence of a flat "plateau" in the curves $i=f(t)$ and $E=f(t)$ at higher Pb concentrations in the catholyte are the subject of the next paper.

DEPARTMENT OF CHEMISTRY AND ELECTROCHEMISTRY, ACADEMY OF MINING AND METALLURGY, CRACOW

REFERENCES

- [1] I. Plaksin, N. Suvorovskaya, IAN SSSR, OTN, (1949), 407.
- [2] A. Rotinyan, B. Drozdov, Zhurn. Obshtch. Khim., **19** (1949), 1843.
- [3] B. Drozdov, Zhurn. Prikl. Khim., **22** (1949), 483.
- [4] A. Bulakh, R. Dratchevskaya, Zhurn. Prikl. Khim., **26** (1953), 1225.
- [5] R. Piontelli, Instituto Lombardi di Science e Lettere d. Sci. **85** (1952), 1.
- [6] R. Glicksman, H. Mouquin, C. King, Journ. of the Electrochem. Soc., **100** (1953), 580.
- [7] S. Krzyżański, Z. anorg. allgem. Chem. **225** (1953), 151; C. A., **30** (1936), 939
- [8] J. Kamecki, J. Sędzimir, Bull. Acad. Polon. Sci. Série A, (1951), 379.
- [9] — Arch. Gór.-Hutn., **3** (1955), 379.
- [10] J. Kamecki, J. Sędzimir, M. Gmytryk, Arch. Gór.-Hutn. **1** (1956), 195.

Kinetics and Mechanism of Cementation Processes.

II. Effect of Cementation Products on Kinetics of the Process

by

J. SĘDZIMIR

Presented by M. ŚMIAŁOWSKI on July 4, 1956

In the present paper it is attempted to give a theoretical explanation of the results obtained when lead was precipitated by zinc in a model cell [1].

1. Theoretical evolution of the kinetic equation

The relation between the rate of the reaction and the current can be found from Faraday's law:

$$(1) \quad -\frac{dm}{dt} = \frac{M}{96500} i,$$

dm denotes the mass of the lead deposited on the cathode in the time dt , and M denotes the gram-equivalent of lead. I assume current efficiency to be 100 per cent since, as was shown by planimetry of curves $i=f(t)$ obtained in experiments discussed in the previous paper [1], the quantity of electricity passed through the solution is equivalent to the quantity of lead precipitated in the course of the process.

Dividing equation (1) by the volume of the reacting solution we obtain:

$$(2) \quad -\frac{dm}{v dt} = -\frac{dc}{dt} = \frac{M}{v \cdot 96500} i.$$

Kinetics of the cementation process are considered predominantly from the point of view of Boguski's equation based on the assumption that diffusion of ions of the metal precipitated through the thin layer of liquid on the surface of the precipitating metal is the slowest stage of the process:

$$(3) \quad -\frac{dc}{dt} = \frac{kF}{v} c$$

k denotes the constant of the reaction rate, F denotes the size of the precipitating surface, v — the volume of the reacting solution, and c — the concentration of the salt of the metal precipitated. In the equation referred to, merely dc/dt and c represent functions of time while the remaining values are assumed to be constant.

I assume that the surface of the metal precipitated is active (ions discharge on the surface of metal already precipitated). This view is confirmed also by Bulakh's observations [2]. The surface grows as precipitation of the metal continues. It ought to be stressed also that the grain of the cementation product grows finer with the decreasing concentration of the solution.

An entire sequence of possibilities of expressing the size of the surface F as a function of time was investigated. Further on, use will be made of the following dependence:

$$(4) \quad F = bF_0 e^{at}.$$

This dependence gives, as shown below, good agreement with experimental data.

The value F from equation (4) was introduced in equation (3); after integration, the following equation was obtained:

$$(5) \quad c = c_0 e^{\frac{kbF_0}{va}(1-e^{at})}.$$

Introducing c from equation (5) in equation (3), we obtain:

$$(6) \quad -\frac{dc}{dt} = \frac{kbF_0 c_0}{v} e^{at + \frac{kbF_0}{va}(1-e^{at})}.$$

Comparing equations (2) and (6) we obtain:

$$(7) \quad i = \frac{96500 kbF_0 c_0}{M} e^{at + \frac{kbF_0}{va}(1-e^{at})}.$$

In order to find the extremum of the curve expressed by equation (7) the derivate di/dt is brought to zero, and the result is

$$(8) \quad t_{mx} = \frac{1}{a} \ln \frac{av}{kbF_0}.$$

Introducing t_{mx} from equation (8) to equation (7), we obtain:

$$(9) \quad i_{mx} = \frac{av}{eM} \frac{96500 c_0}{e} e^{\frac{kbF_0}{va}}.$$

2. Test of the correctness of theoretical assumptions

Equations (8) and (9) may serve to test the correctness of the assumptions underlying the kinetic equation. If these assumptions are correct, then:

$$(10) \quad i_{mx} = \frac{av}{eM} \frac{96500}{e} e^{\frac{kbF_0}{va}} c_0 = A_1 c_0.$$

A_1 , as follows from formula (10), is a function of F_0 , at a constant volume of the solution v .

$$(11) \quad \lg i_{mx} = \lg \frac{av}{eM} \frac{96500}{e} c_0 + \frac{kb}{2 \cdot 303 va} F_0 = A_2 + B_2 F_0.$$

A_2 is a function of c_0 , B_2 is independent of c_0 .

$$(12) \quad t_{mx} = \frac{1}{a} \ln \frac{av}{kb} + \frac{1}{a} \ln \frac{1}{F_0} = A_3 + B_3 \ln \frac{1}{F_0},$$

where A_3 as well as B_3 are independent of c_0 .

Equations (10)–(12) are in agreement with the empirical equations presented in the preceding paper [1]. This appears to confirm the correctness of the assumptions referred to in that paper with regard to the effect of the surface of the metal precipitated on the kinetics of the cementation process.

However, equation (7) and the conclusions which follow from it refer merely to curves with a maximum. The argument presented above fails to explain the presence of a "plateau" in curves $i=f(t)$ at higher concentrations of Pb in the electrolyte.

In order to explain both the causes of this phenomenon, as well as the changes in the potential of the reacting electrodes in the course of the process, it is necessary to consider the curves $i=f(t)$ and $E=f(t)$, obtained in a model cell, in connection with the polarisation curves of the investigated system.

The characteristic curves, $i=f(t)$ and $E=f(t)$, obtained in a model cell [1] are presented in Figs. 1 and 2.

3. Potential curves in dilute solutions (up to 2 g. Pb/l)

As follows from equations (2) and (3),

$$(13) \quad i = AFc$$

(here, A is constant).

The curve in Fig. 1 demonstrates that initially F grows at a higher rate than that at which the concentration decreases; after the maximum is attained the situation is reversed.

A simplified diagram of polarisation curves obtained in the course of the process is presented in Fig. 3

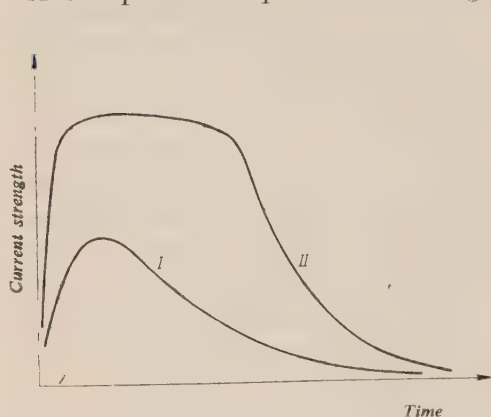


Fig. 1. Current in a model cell as a function of time. Curve I refers to low concentrations (up to 2 g. Pb/l.), curve II to high concentrations (4—6 g. Pb/l.)

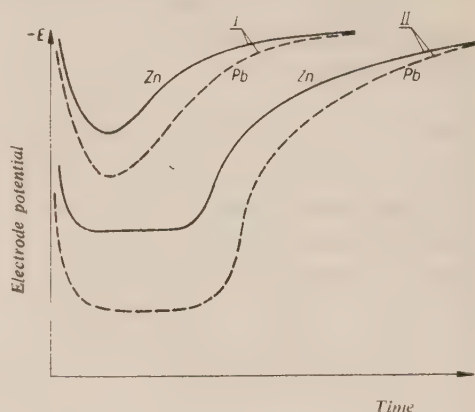


Fig. 2. Electrode potential in a model cell as a function of time. Curves I refer to low concentrations (up to 2 g. Pb/l.), curves II to high concentrations (4—6 g. Pb/l.)

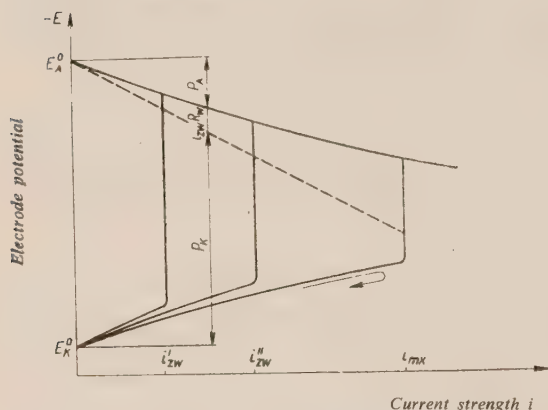


Fig. 3. Shape of polarisation curves in the investigated cell.
(The curves refer to consecutive phases of the process)

The difference in the potentials of the working electrodes, $i_{zw} R_w$, grows linearly with the current and diminishes to zero after the maximum value has been passed. Thus the experimentally established shape of the curves in Fig. 2 is explained.

4. Curves $i=f(t)$ and $E=f(t)$ in concentrated solutions (4—6 g Pb/l)

In order to explain the causes of the presence of "plateaus" on curves $i=f(t)$ (curve II — Fig. 1), the course of the polarisation curves in concentrated solutions is considered (Fig. 4).

- E_K^0 — initial cathode potential,
- E_A^0 — initial anode potential,
- P_A — anodic polarisation,
- P_K — cathodic polarisation,
- i_{zw} — short circuit current (at various intervals of the precipitation process),
- R_w — internal resistance of the cell,
- $i_{zw} R_w$ — potential differences between working electrodes,
- i_{mx} — maximum current (see Fig. 1).

Initially, there flows the current i' conditioned by the extent of the cathodic surface and by the concentration of lead ions in the solution. Increase of the cathodic surface results in the already known increase of current i'' , i''' . The symbol i_{mx} denotes in Fig. 4 the current expressed by equation (9).

At any moment of the process the condition

$$(14) \quad \Delta E_0 = P_K + P_A + iR_w$$

(at $R_z = 0$) must be fulfilled. As follows from both, equation (14) and Fig. 4, the current may grow in this system merely to the value i_p ("pla-

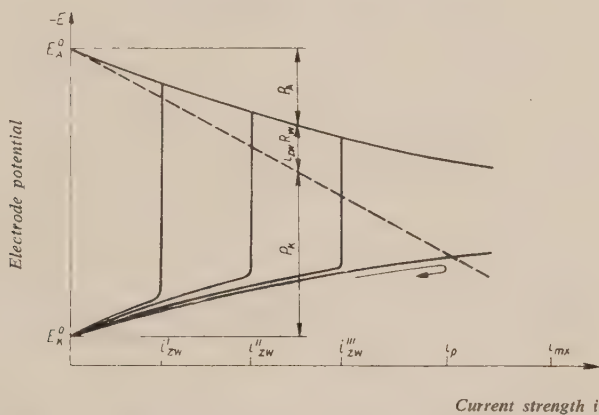


Fig. 4. Shape of polarisation curves in the course of the process in concentrated solutions

- P_K — cathodic polarisation,
- P_A — anodic polarisation,
- iR_w — potential drop on internal resistance,
- i_p — plateau current.

teau" current); further increase is impossible since then $P_K + P_A + iR_w > \Delta E_0$. After i_p has been attained the process proceeds at a constant current value till the concentration of ions of lead becomes so low as to cause a drop of the current, in spite of further growth of the surface. This explains the shape of the curves II in Figs. 1 and 2. This reasoning was confirmed in the following experiment: two precipitation processes were run in solutions containing 2 g. Pb/l., (a) by using a 30-sq. cm. F_{Zn} anode, and (b) by using a 3-sq. cm. F_{Zn} anode. As was expected, the curve $i = f(t)$ shows a maximum in the first case and a "plateau" in the second (using a smaller anodic surface we obtain a steeper curve of anodic polarisation, and, consequently, i_p can be attained). It should also be mentioned that the decrease of current oscillations within the "plateau" [1] testifies to a change of the factor controlling the rate of the process.

5. Conclusions

The cementation process mechanism quoted above, advanced on the basis of experiments conducted in a model cell, appears to explain the discrepancies encountered in literature, relating to precipitation of metals by metals.

Depending on the texture of the precipitation product, cementation processes can be divided into three basic groups:

a. Loose deposit, with a poorly developed surface — cathodic control; the process more or less follows Boguski's equation. The constant " k " is greater than could be expected from diffusion data.

b. Loose deposit, with a well developed surface — cathodic control; the constant " k " calculated with the aid of Boguski's equation grows in the course of the process (or shows notable discrepancy).

Groups a. and b. account for the observations quoted in [2] to [6].

c. A solid layer of cementation products covers the surface of the precipitating metal. Anodic polarisation curve is steep, i_p is easily attained. From this moment on, the reaction proceeds according to zero order kinetic equation. The precipitating metal may be also covered with a tight layer of the metal precipitated. This corresponds to the observations quoted in papers by Rotinyan [7], Drozdov [8] and by Krzyżański [9].

DEPARTMENT OF CHEMISTRY AND ELECTROCHEMISTRY, ACADEMY OF MINING AND METALLURGY, CRACOW

REFERENCES

- [1] J. Sędzimir, M. Pawełkowa, in this issue p. 719.
- [2] A. Bulakh, R. Dratchevskaya, Zhurn. Prikl. Khim. **26** (1953), 1225.
- [3] I. Plaksin, N. Suvorovskaya, IAN SSSR, OTN, (1949), 407.
- [4] R. Glicksman, H. Mouquin, C. King, Journ. of the Electrochem. Soc. **100** (1953), 580.
- [5] J. Kamecki, J. Sędzimir, Bull. Acad. Polon. Sci. Série A (1951), 379.
- [6] J. Kamecki, J. Sędzimir, M. Gmytryk, Arch. Gór.-Hutn. **1** (1956) 195.
- [7] A. Rotinyan, B. Drozdov, Zhurn. Obshtch. Khim. **19** (1949), 1843.
- [8] B. Drozdov, Zhurn. Prikl. Khim. **22** (1949), 483.
- [9] S. Krzyżański, Z. anorg. allgem. Chem. **225** (1953), 151; C. A., **30** (1936), 939.

Sedimentary Characteristics of the Jarmuta Beds (Maestrichian) in the Pieniny Klippen Belt (Central Carpathians)

by

K. BIRKENMAJER

Presented by M. KSIĄŻKIEWICZ on July 11, 1956

Position of the Jarmuta beds in the stratigraphic profile of the Klippen mantle

The denotation "Jarmuta sandstone" was introduced by L. Horwitz [4] to define the conglomerate-sandstone series from the vicinity of Szczawnica, Szlachtowa and Jaworki in the Pieniny Klippen belt. On the basis of foraminifera the age of the Jarmuta sandstones was defined by P. Bieda [1] as Maestrichian.

The Maestrichian age of the Jarmuta beds established on the basis of large foraminifera identified by F. Bieda must now be reconsidered, as A. Papp [10] found *Siderolites vidali* also in the Campanian and *Orbitoides media* s. l. both in the Campanian and Maestrichian. It is thus possible that the Jarmuta beds represent both the Campanian and Maestrichian.

The Jarmuta beds can be divided into several facies: 1) the deepest facies is of the Flysch type; 2) the conglomerate-sandstone facies represents a more shallow deposit; 3) a littoral facies with features characteristic for cliff deposits; 4) finally conglomerate deposits of the supralittoral type. The presence of cliff and supralittoral deposits marks in the Klippen belt the south coast line of the Upper Cretaceous marine basin. To the south of this line extended at that time the unsubmerged parts of the Middle Cretaceous orogene formed of the Branisko and Pieniny series which were unaffected by transgression till the Lower Eocene stage [2].

It follows, therefore, that the Jarmuta beds represent an example, rather uncommon for the Carpathians, of Flysch deposits with a definite relation to a known coast of a sedimentation basin.

Over the Jarmuta sediments lie continuously variegated beds composed of cherry-red and variegated marly-clayey shales with thin intercalations of green sandstones. Subsequently (in the Paleocene) the Laramide orogeny caused the folding of the Jarmuta and variegated beds in the Klippen belt and this was followed in the Lower Eocene by the Magura Flysch transgression which covered the still unsubmerged part of the Middle Cretaceous orogene.

The supralittoral facies

The sediments of this facies consist of conglomerates or gravels consolidated poorly or not at all and are composed of well and very well-rounded boulders. They contain no fossil fauna from which their age could be established directly. The diameter of approximately 80 per cent of the gravel material is 1-6 cm., in 5-10 per cent of the material the diameter is 10-40 cm. and the remainder consists of grains up to 1 cm. in diameter and of the cementing medium. The boulders are arranged in irregular feebly marked beds or even absolutely at random. The thickness of these sediments is at the most about 50 m.

The gravels are composed exclusively or almost so of material exotic in respect to the Klippen belt.

In this material Middle Triassic limestones and dolomites predominate making up about 40-50 per cent of the total, then come Werfenian quartzites corresponding to about 40 per cent. The other sedimentary rocks, metamorphic and volcanic rocks, and the most rarely encountered plutonic rocks amount together to about 5-15 per cent. The material of the Klippen series does not exceed 5 per cent and ranges from well or very well-rounded exotic boulders through scarcely rounded fragments to sharp edged angular fragments. It is composed of red and green radiolarites of Lower Malm, greyish cherty Neocomian limestones of the Branisko series, red Bathonian crinoidal limestones, nodular limestone (Callovian-Kimmeridgian) and red Calpionella Tithonian limestones of the Czorsztyn series.

In view of the above features the gravels and conglomerates may be considered to be of river origin and were deposited in the supralittoral zone in the form of alluvial cones. The remarkably high proportion of rounded exotic limestone and dolomite boulders may indicate that they were transported over a relatively short distance by rapidly flowing water. Evidence that the material was deposited in a zone neighbouring with a marine coast-line may be supplied by the fact that in the vicinity of the village of Dursztyn marine cliff sediments of the Jarmuta beds were deposited on these gravels.

Cliff facies

The deposits belonging to this facies appear discontinuously at the base of the southernmost tectonic scales of the Jarmuta beds. Their presence in the Klippen belt has been established over a distance of about 30 kilometres between the villages of Krempachy and Biala Woda. The thickness of these deposits is relatively small and does not exceed 2 to 5 m. They usually consist of limestone breccias and conglomerates strongly cemented by a calcareous matrix and composed of local material from the Klippen series that changes from place to place according to the geological structure of the primarily existing coast.

The material from the Klippen series is mostly unrounded or rounded only very slightly. Its fragments are either very small, not exceeding several millimetres and giving after cementation detritic limestones or hornstones, or more frequently the diameter of the fragments is on the average 2-4 cm., or finally in some few instances the diameter may be even 0.5-0.7-1.0 cm. In some outcrops only one component predominates, e. g. cherty limestones or radiolarites, whereas elsewhere the predominant element may be almost exclusively crinoidal limestone or nodular limestone.

Exotic materials are either entirely absent or their proportion is only 1-5 per cent. They consist of well or very well-rounded fragments of quartz and macroscopically undistinguishable magmatic and sedimentary rocks together with irregular fragments of crystalline shales ranging in size from 0.5 to 2-4 cm.

The matrix of the cliff breccias and conglomerates consists of re-crystallised calcareous pelite containing debris of inoceramus and oyster shells and sometimes too broken and rounded fragments of belemnites and segments of crinoids washed out from older rocks.

The predominance of the harder local material points to selection and desintegration of the shaly-marly sediments by surf waves. As the exotic material is highly rounded it seems possible that it was transported to the littoral zone by rivers from the inland.

The bedding of the cliff sediments is, in irregular beds, several to several tens of centimetres thick. The larger flat rock fragments lie conformably with the bedding. The material is either unsegregated in respect to size or of similar size. There are no traces of graded bedding.

The substratum of the cliff sediments consists of the Branisko series. The cliff beds were deposited on that series in the course of a progressive marine submersion when favourable conditions for their preservation existed. Most probably it is a facies spreading diachronously to the south, as in several cases cliff deposits were observed to be overlying sediments of the supralittoral type. The breadth of this submersible zone probably did not exceed several hundred meters.

Conglomerate-sandstone facies

The sediments of this facies overlie the cliff facies to the south, while to the north they lie directly on the Branisko series. The breadth of the present-day zone occupied by the conglomerate-sandstone facies is approximately 500 m. but previously, before faulting occurred, the zone was possibly from several to twenty kilometres wide. In the southern part, the thickness of the facies is about 30 m. but to the north it increases to 150-200 m.

Sandstones are more common in these beds than conglomerates. The sandstones are fine, medium- and coarse-grained and the conglomerates are fine- or medium-grained, and rarely coarse-grained. The conglomerates form either the underside of the beds or intercalations in the sandstones several centimetres to several metres thick. The thickness of the sandstone beds ranges from several centimetres to more than one metre. There usually are no shales or they appear in very small quantities, the ratios of shales to sandstones being 1:10, 1:20 and even 1:30.

Of the bedding structures occurring in sandstone and conglomerate beds the graded type is most common, about $\frac{2}{3}$ of all the bedding structures being of this type. Laminated bedding is also very frequent. These beddings occur either singly or just as often as composite structures. Undulated and slip beddings are quite exceptional and laminated and nonlaminated current (diagonal) beddings have never been observed.

In graded layers the coarser grains may be assembled in the lower layers, the grains becoming finer towards the top. This is the normally graded bedding (Fig. 5 the lower part of the layer). The grading may be complete when the distribution into grades is gradual, or incomplete *) when some grades are missing in the vertical sequence. In all cases, the grade characteristic for some given layer has also a smaller proportion of grains of finer or coarser grades. This effect has been already reported by Ph. H. Kuenen [7], M. Książkiewicz [6]. Inversed graded bedding, i. e. when the grains are finest at the bottom and become coarser towards the top, is less frequent. On the other hand, more common are symmetrically inversed graded beddings in which a layer begins at the bottom with fine grains becoming coarser to the top and then again getting finer. Usually the symmetry is not complete, but according to the definition introduced by M. Książkiewicz [6] *pensymmetric* (Fig. 11). The layers with symmetrically inversed beddings may be repeated several times (Figs. 7, 10). Quite frequent too are beds or layers of sandstone built of grains of one size (non-graded bedding) which in the Flysch stage are described by S. Dzułyński and

*) This latter type is defined in [6] as "discontinuous".

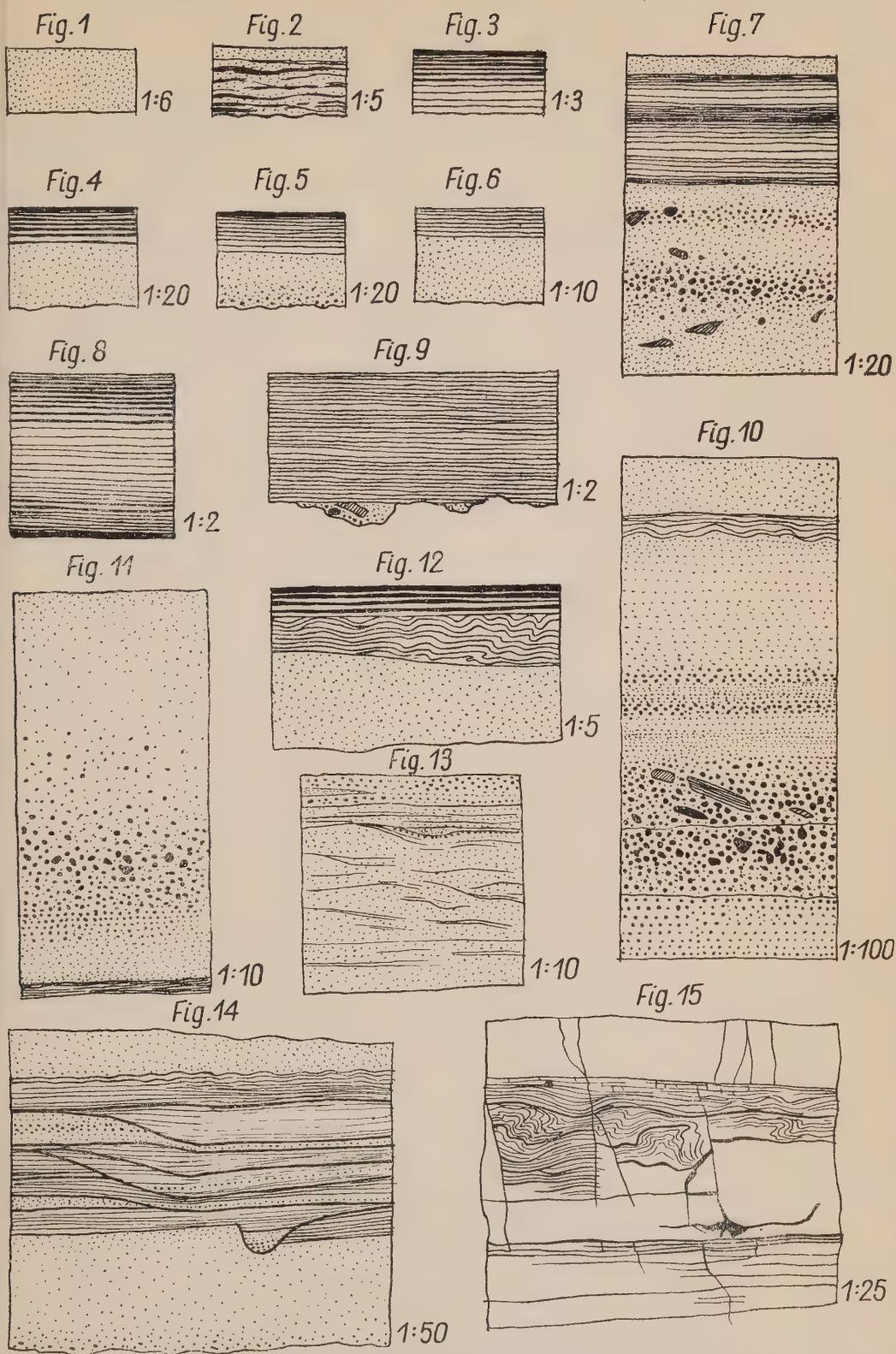


Plate I. The more common types of bedding in the sandstones of Jarmuta beds. For more detailed explanations see text.

A. Radomski [3] as "homogenous sandstone". Non-graded beddings may form independent beds (Fig. 1) or they may occur either at the bottom (Figs. 4, 6, 12) or at the top (Figs. 7, 10) of beds with composite structure.

Laminated bedding appear either independently or in the upper parts of a bed with a graded structure. The particular laminae may be of lenticular shape (Figs. 2, 13), but more often they are even. In the latter case the bedding is without gradation and is evenly (Fig. 6) or variably laminated. Instances are however, very common of laminated beddings which exhibit either normal symmetrical gradation (Figs. 4, 5, 13), or single or multiple (Fig. 7) inversed symmetrical gradation.

Plant and mica detritus is common in the sandstones and is usually accumulated in beds with laminated structure. Moreover, quite exceptionally large foraminifera from the genera *Orbitoides*, *Lepidorbitoides* and *Siderolites*, and fragments of calcareous algae are also found. More frequent are fragments of inoceramids and oyster shells which occur at the bottom of layers with graded structure where the grains are of the same size as these fragments.

Slip and undulated bedding structures are less common than the two types already described and usually occur in the middle or at the top parts of layers (Figs. 10, 12). The slip movements had NNE and NNW directions. Flow markings appearing on the under side of the sandstone beds where they contact with shales are fairly frequent and display, too, a similar direction of flow towards NNE. Very common are traces of submarine erosion caused by turbidity currents (Fig. 14) of the type similar to that described by Ph. H. Kuenen ([7], Fig. 14; [8], Fig. 1; [9], Fig. 1H, 14), and by M. Książkiewicz (see [6] Fig. 9). The erosion channels run in roughly one direction approximately towards NNE. The absence of shales and the circumstance that the sandstone layers lie directly one on the other, frequently on a distinctly stripped bedding surface (Fig. 14), seem to indicate an intense erosion by turbidity currents of considerable energy. Because of this the prevailing conditions did not favour the preservation of groove marks, which according to S. Dżułyński and A. Radomski [3] are made by shale fragments carried in the lower parts of a turbidity current thus grooving a clayey bottom.

The material of the conglomerate interbeddings and intercalations is predominantly exotic in nature. The exotic elements are in this case 2-5-10 cm. in diameter and are well or very well-rounded. Crystalline material is more common than that of sedimentary origin.

Overlying the conglomerate-sandstone series is the Flysch series. The Flysch series is from about 20 m. to about 50 m. thick. Above it are the variegated beds which were up to 100 m. thick.

The Flysch facies

The beds of this facies occur mainly in the part of the sedimentation Jarmuta basin lying furthest to the north where they form a zone of tectonic shells approximately 1 km. wide. The relation to the substratum of the Klippen series is unknown. In the south, the Flysch facies overlies the conglomerate-sandstone facies.

Originally the exclusively Flysch zone of the Jarmuta beds was at least several and perhaps even nearly 20 km. wide. To the south, where the Flysch sediments overlie the Kliff breccias and the conglomerate-sandstone facies, their thickness is about 20 m. but it increases up to about 500 m. to the north.

The sandstones of the Flysch facies are fine-grained, their beds are only from several to 20 cm. and rarely to 1 m. thick and contain interbeddings of marly shales of about the same thickness. The ratio of shales to sandstones varies from 1:1 to 2:1.

The most commonly observed bedding structure is of the graded type, but laminated bedding is almost equally frequent. In respect to structure the beds are homogenous or composite. Among the graded bedding structures there are those with a normal complete or incomplete sequence (Fig. 5) and non-graded ones (Figs. 1, 4). Among laminated bedding there are instances of even laminations without gradation (Fig. 9), varied laminations, and also, quite frequently, of normal (Fig. 3) or inversed asymmetrical gradations. The last mentioned bedding type may occur independently, it may form the upper part of beds which in their lower layers have a graded bedding, or it may form the middle layers of beds which are graded below and above. There are some instances of inversed, symmetrically laminated bedding (Fig. 8). Frequently lamination is marked by plant and mica detritus.

Alongside of the types already described, lenticular laminated beddings are also observed. The effect similar to that reported by M. Książkiewicz [5] and defined by him as slip bedding and by Ph. H. Kuenen (see [7], Fig. 2, [8], Fig. 10, [9], Fig. 5) as slump structure is here more frequent than in the previous facies. Moreover, in the arrangement of the laminae there are disturbances similar to those reported by J. L. Rich [11] from the Silurian rocks of Wales and defined as intrastratal contortions and by Ph. H. Kuenen (see [7], Fig. 2); [8], [9] as convolute bedding. The direction of both the slip and convolute types of bedding is predominantly towards NNE. Diagonal (current) bedding is relatively rare.

In the Flysch facies markings on the under side of the beds are rather common. Sometimes the markings are of organic origin, but most often of mechanic origin. Among the latter type the most common are flow markings directed towards NNW and NNE as well as with tendencies approaching directions towards ENE or WSW. Grooves are

comparatively rare, possibly because of the absence in the sandstones of shale fragments which could make them. Traces of submarine erosion and of washing out are rare.

Intercalations of fine and medium-grained conglomerates are seldom observed in the Flysch facies. The material of which these conglomerates are composed is here similar to that in the facies described previously.

LABORATORY OF GEOLOGY AND STRATIGRAPHY, POLISH ACADEMY OF SCIENCES

REFERENCES

- [1] F. Bieda, *Compte rendu des recherches sur les Foraminifères du Flysch de la zone piénine des Klippes. C=R.* Inst. Geol. Pol. (1935), No. 42.
- [2] K. Birkenmajer, *On the age of the so-called "Puchov marls" in the Pieniny (Central Carpathians) and stratigraphy of the Pieniny Klippen-belt mantle.* Inst. Geol. Biul. (1954), No 88.
- [3] St. Dżułyński, A. Radomski, *Origin of groove casts in the light of turbidity currents hypothesis.* Acta geol. Pol. **5** (1955).
- [4] L. Horwitz, *Études stratigraphiques sur la couverture des Klippes piénines (Karpates Polonaises). I-ère partie.* Inst. Geol. Pol. Bull. (1938), No. 9/2.
- [5] M. Książkiewicz, *Slip bedding in the Carpathian Flysch.* Ann. Soc. Geol. Pol., **19** (1950).
- [6] — *Graded and laminated bedding in the Carpathian Flysch.* Ann. Soc. Geol. Pol., **22** (1954).
- [7] Ph. H. Kuenen, *Palaeogeographic significance of graded bedding and associated features.* Koninkl. Nederl. Akad. Wetensch. Proc. s. B, **55** (1952).
- [8] — *Graded bedding with observations on Lower Paleozoic rocks of Britain.* Verh. Koninkl. Nederl. Akad. Wetensch., Aft. Nat., Eerste Reeks, **20** (1953), No. 3.
- [9] — *Significant features of graded bedding.* Bull. Amer. Ass. Petrol. Geol., **37** (1953), No. 5.
- [10] A. Papp, *Über die Entwicklung von Pseudorbitoides und Lepidorbitoides in Europa.* Verh. geol. Bundesanst. (1954), No. 3.
- [11] J. L. Rich, *Flow markings, groovings, and intrastratal crumplings as criteria for recognition of slope deposits, with illustrations from Silurian rocks of Wales.* Bull. Amer. Ass. Petrol. Geol., **34** (1950), No. 4.

